

**RELIABILITY AND VALIDITY OF THE PREMIE-NEURO IN DETECTING
EARLY NEURODEVELOPMENTAL DELAY AND DISABILITY IN
PRETERM INFANTS**

Kendra L. Gagnon
MPT, University of Missouri-Columbia
BHS, University of Missouri-Columbia

Submitted to the graduate degree program in Rehabilitation Science and the Graduate
Faculty of the University of Kansas in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

Lisa A. Stehno-Bittel, PT, PhD
(Chair)

Byron J. Gajewski, PhD

Patricia S. Pohl, PT, PhD

Jeff D. Radel, PhD

Kathleen B. Weatherstone, MD

July 24, 2009
Date of dissertation defense

The Dissertation Committee for Kendra L. Gagnon certifies that this is the approved version of the following dissertation:

**RELIABILITY AND VALIDITY OF THE PREMIE-NEURO IN DETECTING
EARLY NEURODEVELOPMENTAL DELAY AND DISABILITY IN
PRETERM INFANTS**

Committee:

Lisa A. Stehno-Bittel, PT, PhD
(Chair)

Byron J. Gajewski, PhD

Patricia S. Pohl, PT, PhD

Jeff D. Radel, PhD

Kathleen B. Weatherstone, MD

July 24, 2009
Date of dissertation defense

Abstract

A disproportionate amount of dollars are spent to provide specialized early intervention services to a number of preterm children in the first months and years of life, despite mounting evidence that these services benefit only a small percentage of preterm babies who are truly at high risk for neurodevelopmental disability and delay. Unfortunately, there is no “gold standard” test or procedure for identification of at-risk preterm infants, and it is extremely difficult for the neonatal care provider to determine which preterm babies will gain the most benefit from early intervention services. The Premie-Neuro is a new standardized clinical neurological assessment tool that may aid the clinician in identifying at-risk preterm infants, but its psychometric properties have not been established. In this study, we set out to answer three primary questions about the Premie-Neuro. Does the Premie-Neuro have adequate reliability and construct validity for clinical use? Is Premie-Neuro performance predictive of neurodevelopmental disability and delay at term and post-term ages? Are there modifications of the Premie-Neuro that may enhance its reliability, validity, predictive value, and clinical use?

In Chapter 2, we evaluated the reliability and construct validity of the Premie-Neuro by using the tool to test 34 preterm infants in the neonatal intensive care unit (NICU). The Premie-Neuro was administered to each infant twice by the same examiner, no more than 72 hours apart, then continued to be administered by the same examiner weekly through 37 weeks post-menstrual age (PMA) or discharge from the NICU. One Premie-Neuro assessment per infant was observed and scored

by a second examiner. At NICU discharge, infants' medical charts were reviewed to determine risk for poor neurological outcomes. Our results showed that the Premie-Neuro raw scores had fair to moderate test-retest and interrater reliability and discriminated between groups of infants known to differ in terms of risk. However, the Premie-Neuro classifications (abnormal, questionable, or normal) based on the authors' recommended raw score cut-points were not reliable or valid. This study showed that the Premie-Neuro raw scores had acceptable reliability and validity for use by the clinician to identify at-risk preterm infants, but that the Premie-Neuro classifications should be interpreted cautiously.

Predictive validity of standardized assessment tools is of great interest to the neonatal care provider. A predictive tool allows the clinician to make a long-term prognosis for a preterm infant's development and function, and can provide guidance for discharge planning and follow-up care. In Chapter 3, we examined whether Premie-Neuro performance in the NICU was predictive of outcomes at term and three months adjusted age. As in Chapter 2, we tested preterm babies using the Premie-Neuro each week through 37 weeks PMA or NICU discharge. Infants were then tested at term adjusted age (38-42 weeks PMA) using the NeoNeuro and at 3 months adjusted age using the Alberta Infant Motor Scale (AIMS) and Infanib. We found that, while early Premie-Neuro scores had some predictive value, Premie-Neuro raw scores obtained just before NICU discharge were significantly predictive of performance on all assessments administered at term and three months adjusted age. However, the Premie-Neuro classifications were generally not predictive of term and

three-month outcomes. We concluded that the clinician may use discharge Premie-Neuro raw scores to predict term and three-month outcomes for preterm infants. Again, we recommended that the clinician use caution when interpreting Premie-Neuro classifications based on the authors' recommended raw score cut-points.

In Chapters 2 and 3, we found the Premie-Neuro raw scores—but not classifications—to be psychometrically sound. Thus, in Chapter 4, we set out to determine whether modifying the raw score cut-points for making Premie-Neuro classifications may improve the tool's reliability and validity. In Chapter 4, we retrospectively analyzed predictive validity data presented in Chapter 3 to determine whether infants with normal versus abnormal outcomes at term and/or three-months adjusted age performed significantly differently on the Premie-Neuro in the NICU. We found that, after 34 weeks PMA, mean Premie-Neuro scores for infants with abnormal follow-up were less than 90, and that infants with discharge Premie-Neuro raw scores less than 90 performed significantly worse on term and 3-month follow-up assessments than infants with scores 100 or greater. Using a discharge Premie-Neuro raw score cut-point of <90 (rather than the authors' recommended cut-point of <100) to classify a Premie-Neuro assessment as not normal improved reliability, construct validity, and predictive validity of the Premie-Neuro classifications. We concluded that a raw score cut-point of <90 may be used by the clinician to identify at-risk infants after 34 weeks PMA, but cautioned that this strict cut-point may result in a handful false-negatives. Thus, preterm infants with borderline raw scores, ranging

from 90-99, should be closely monitored to ensure that neurological abnormalities do not emerge in those babies.

In summary, this body of work is the first to show that the psychometric properties of the Premie-Neuro raw scores meet reliability and validity criteria for clinical use in the fragile, highly variable preterm infant population. Although Premie-Neuro classifications based on raw score cut-points recommended by the assessment's authors may not be reliable and valid, the present study proposes a new set of cut-points that may be used by the clinician to reliably classify a preterm infant's assessment as normal, abnormal, or questionable. These findings may have important implications for clinical practice in the NICU, as we have shown that clinicians now have access to a psychometrically sound assessment that may be administered repeatedly throughout an infant's NICU stay—the Premie-Neuro. Our proposed guidelines for interpreting Premie-Neuro assessments as normal, abnormal, or questionable may be used by clinicians to identify at-risk babies who will benefit from specialized early intervention services and/or close post-term follow-up. This has potential to improve effectiveness, efficiency, and outcomes of early intervention services provided to preterm babies.

Acknowledgements

First and foremost, I would like to thank my research mentor and committee chair, Dr. Lisa Stehno-Bittel, whom I first met almost six years ago when I was a young PT who simply had an interest in teaching. It was during our first meeting that Dr. Bittel encouraged me to apply to the PhD program at KUMC. Somehow, she made the idea of conducting my own research—something that had once been unthinkable to me—seem possible. Two years later, I began this journey. And after four years of Dr. Bittel's advice, support, encouragement, and optimism, I have completed a research project of which I am proud. Thank you, Dr. Bittel.

I would also like to thank the members of my committee, past and present. Dr. Steven Barlow, Dr. Byron Gajewski, Dr. Patricia Kluding, Dr. Patricia Pohl, Dr. Jeff Radel, and Dr. Kathleen Weatherstone have all contributed their time and expertise to my doctoral studies and this project. As I read this dissertation, I see the unique contributions each has made to my learning and this project, and I am fortunate to have worked with each of them along the way.

I would like to thank Dr. Katherine Schooley, Dr. Robert Holcomb, Dr. Christopher Stapley, and Dr. Thomas Lancaster of Sunflower Neonatology; Dr. Donna Daily of Vanderbilt University; and Susan Cannon of the University of Kansas Medical Center. I could not have completed this project without their support or the support of the NICU nurses and staff—particularly Jessie Charpentier and Susan Ise—at Overland Park Regional Medical Center. Thank you to the families of the Overland Park Regional Medical Center NICU who allowed their babies to

participate in this study in hopes that future babies may receive better care. As a mother myself, I know how much these families trusted in me, and I was humbled by that trust with each baby and family I met during the study.

I am grateful to the PhD students, faculty, and staff in the Department of Physical Therapy and Rehabilitation Science at the University of Kansas Medical Center for their ongoing support and camaraderie. Thank you to Dr. Irina Smirnova, the director of the PhD program, for periodically “checking in” and always providing help and guidance when needed. And to the “Taco Tuesday” crew—Katie, Sandy, Eric, and Steve—thank you for swapping stories, sharing experiences, and giving me an outlet when I just needed to vent.

Finally, I would like to thank my family. To my parents, who instilled in me a love of learning from an early age. To my son Max—who was born during my work on this project—for making me laugh, forcing me to keep things in perspective, and teaching me something about myself every day. To my soon-to-be-born little girl—as I write this, I do not forget that many of the babies in this study were born at the gestational age you are today. I am thankful for your health and hope that this work may help you if you, too, are born too soon. And most importantly, I thank my husband and best friend, Joe. When I wondered if I could do this, you said there was no way I couldn’t. When I struggled with balancing work and family, you reminded me that the best thing I could give Max was a mother who was smart and strong and fulfilled. I have never known anyone else like you, and every day I am thankful that you chose me as your partner. I dedicate this dissertation to you.

Table of Contents

Abstract.....	iii
Acknowledgements	vii
Table of Contents	ix
Chapter 1	1
1.1 Overview	2
1.2 Prematurity	4
1.3 Neuromaturation and mechanisms for brain injury in the preterm infant	5
1.4 Neurodevelopmental outcomes in preterm infants.....	9
1.5 Predictors and risk factors for poor neurodevelopmental outcomes.....	11
1.6 Effects of early intervention.....	13
1.7 The role of clinical neurological assessment in identification of at-risk	
infants.....	16
1.8 The Premie-Neuro.....	19
1.9 Significance of presented work	20
1.10 Specific aims and statement of hypotheses	23
1.11 Summary.....	25
Chapter 2	27
2.1 Abstract.....	28
2.2 Introduction.....	30
2.3 Methods.....	33
2.4 Results	36

2.5 Discussion.....	39
2.6 Conclusions.....	44
2.7 Figure Legend.....	53
Chapter 3	56
3.1 Abstract.....	57
3.2 Introduction.....	59
3.3 Methods.....	61
3.4 Results	68
3.5 Discussion.....	70
3.6 Conclusions.....	74
3.7 Figure Legend.....	81
Chapter 4	85
4.1 Abstract.....	86
4.2 Introduction.....	88
4.3 Methods.....	90
4.4 Results	95
4.5 Discussion.....	99
4.6 Conclusions.....	103
4.7 Figure Legend.....	111
Chapter 5	118
5.1 Summary of findings.....	119
5.2 Clinical implications	123

5.3 Limitations.....	129
5.4 Future directions.....	135
5.5 Conclusions.....	140
References.....	142

Chapter 1

Introduction

1.1 Overview

The United States Department of Health and Human Services reports that, in 2006, 12.8% of all infants born in the United States were born preterm. The rate of preterm births—defined as birth prior to 37 completed weeks gestational age (GA)—has increased 20% since 1990 and 36% since 1980 (1). The cause of this increase is unknown. Prematurity is one of the leading causes of infant death and major disabilities such as cerebral palsy and mental retardation (1, 2). Advances in neonatal care, including the introduction of antenatal steroids and postnatal surfactant, led to improved survival rates among preterm infants through the 1990s. However, this increase in survival rate was accompanied by an increase in the number of preterm infants with major medical complication and poor neurodevelopmental outcomes (2-5). Over the last decade, continued improvements in medical care have led to improved outcomes of preterm infants, and the proportion of preterm infants with neurodevelopmental deficits, although not declining, has not risen (6, 7). However, due to the rising numbers of preterm birth and improved awareness and methods of detection, there continues to be increasing numbers of preterm children with identified neurodevelopmental disability and/or delay. During their stay in the neonatal intensive care unit (NICU) and in the first years of life, preterm children are often provided early intervention services such as speech, physical, or occupational therapy. However, research is unclear whether these types of early intervention services are effective in improving long-term outcomes—particularly neuromotor outcomes—in preterm children who receive such services simply because they are

preterm (8-11). Some studies suggest that motor interventions are effective in improving outcomes when they are aimed only at those preterm infants who are identified as high-risk using a standardized assessment (12, 13). Thus, there is a growing need for clinical neurological assessment tools that will allow health care providers to reliably identify preterm infants most at-risk for neurological abnormality and potential disability so that those infants may be selectively referred for early intervention physical therapy services, avoiding over-utilization of services by preterm children who are not at-risk.

Currently, there are a handful of clinical neurological assessment tools available for use with preterm infants in the NICU, but there is no agreed-upon “gold standard” for neonatal neurological evaluation. Available assessments can be lengthy to administer, require hours, weeks, or even months of administrator training, and/or cannot be given to children less than 30-32 weeks post-menstrual age. This makes them very impractical for use by clinicians who desire an assessment tool that requires minimal infant handling and can be given quickly and repeatedly from birth to term, even for those infants who are born as early as 23-24 weeks gestational age. The Premie-Neuro—a new standardized neurological assessment tool developed by Donna Daily, MD—is unique in that it requires no special training for neonatal care providers and can be given in approximately 10 minutes to children ages 23 through 37 weeks post-menstrual age (PMA) (14). The work presented in this dissertation will describe the reliability and validity of the Premie-Neuro. In this study, we examine interrater and test-retest reliability of the Premie-Neuro. We also determine

whether the Premie-Neuro can discriminate between groups of infants who differ in terms of risk for neurodevelopmental delay and disability, and whether performance on the Premie-Neuro in the NICU can predict neurological and neuromotor performance at term and 3 months adjusted age. Finally, we suggest raw score cut-points that can be used to identify neurologically abnormal or questionable infants who may benefit from further examination and/or referral for early intervention services. This work has implications for clinical assessment of preterm infants and will impact clinical decision-making when determining which infants require further examination or specialized early intervention services.

1.2 Prematurity

Prematurity is defined as birth prior to 37 weeks completed gestation. The preterm birth rate in the U.S. has been increasing steadily for the last several decades. Since 1990, the rate of moderately preterm (less than 34 weeks gestation) and very preterm (less than 32 weeks gestation) births has risen modestly, but the rate of late preterm birth (34-36 weeks gestation) has increased sharply (1). Late preterm birth now accounts for approximately 60-70% of all preterm births (15). Increasing multiple births has no doubt contributed to the rise in preterm births. However, the preterm birth rate for singletons is also rising—11.1% of singletons were born preterm in 2006—so multiple births cannot be solely responsible for the trend. The rate of preterm birth is highest for black infants (18.5%). The preterm rate for Hispanic births is 12.2% and 11.7% for non-Hispanic white births (1).

Approximately 30% of preterm births are medically indicated, usually due to maternal or fetal infection. 40-45% of preterm births occur following spontaneous preterm labor and 25-30% following premature rupture of membranes. Although there are a number of maternal risk factors associated with preterm birth—including black race, short interpregnancy interval, low pre-pregnancy body mass index, previous preterm delivery, vaginal bleeding, stress and depression, smoking, intrauterine infection, and biological and genetic markers—a precise cause of preterm birth cannot be determined in most cases (15). Until the causes of prematurity are better understood, it is likely that the preterm birth rate will continue to rise.

1.3 Neuromaturation and mechanisms for brain injury in the preterm infant

Neuromaturation—the development of the central nervous system (CNS)—is a dynamic process, dependent not only upon a preprogrammed genetic progression, but also upon the interaction between a fetus or infant and his environment. Genetic processes ensure that early brain development in all humans follows the same predictable sequence. However, intrauterine and extrauterine experiences result in individual variability in the developing brain (16, 17).

In the first several weeks of gestation the neural plate—which eventually gives rise to the central nervous system—is formed. Shortly after, neuroblast proliferation and neuronal migration, aggregation, and differentiation begin. At approximately 8 weeks gestation, the first layer of the cortical plate is in place and the first synapses are seen (17, 18). It is about this time that the earliest primitive

reflexes—such as rooting and grasp—emerge (16). Neurons continue to migrate and aggregate, forming deeper cortical layers as gestation progresses. By 8 months GA, cortical layers resemble that of older children. During the third trimester of pregnancy—as the neurons reach their final positions—dendritic growth and synapse formation accelerate as each cortical neuron forms approximately 1,000 synaptic connections (19). From 28 weeks PMA, synaptic density increases as much as 6 times until it peaks as early as 3 months post-term in sensory areas such as the visual and auditory cortices and as late as 15 months post-term in the prefrontal cortex (18).

In the second half of gestation, beginning at approximately 24 weeks GA, myelination of the central nervous system begins (17). Myelination of the lower, subcortical motor system progresses upward from the spinal cord between approximately 24 and 32 weeks GA. During this time, reflexes become stronger and flexor tone develops in a caudocephalic direction. At 32 weeks GA, myelination of the higher corticospinal pathways begins in the pons, progressing down to the spinal cord and up to the cortex. Development of these higher motor pathways suppresses the lower motor system, resulting in a decrease in the dominance of involuntary reflexes and an increase in head and trunk control and voluntary movements (16). Myelination continues after birth—reaching peak velocity between birth and one year as infants develop postural reactions and fine motor control—and persists well beyond the 3rd decade of life (16-18).

Ironically, although normal rapid myelination occurring around the time of birth is necessary for healthy brain development, it also makes the neonatal brain

particularly susceptible to injury. Oligodendrocytes—the cells responsible for myelination in the CNS—are extremely vulnerable to hypoxia (17, 19). This is more problematic for preterm infants, the youngest of whom may be born just as myelination is beginning, because immature oligodendroglial precursor cells are especially susceptible to injury. This problem is compounded by the fact that, in the preterm infant, blood supply to cerebral white matter is incomplete and regulation of cerebral blood flow is impaired (17, 19, 20). Thus, damage to the white matter around the ventricles of the brain is common in preterm infants, an injury known as periventricular leukomalacia (PVL). PVL is the primary form of brain injury in the preterm infant and is associated with symmetrical damage to the internal capsule (20, 21). This results in the most common form of cerebral palsy in preterm children—spastic diplegia—in which spasticity is present in all four limbs but is dominant in the lower extremities (3).

The other major form of brain injury frequently associated with preterm birth—and one that is frequently associated with PVL (22)—is periventricular hemorrhage/intraventricular hemorrhage (PVH/IVH). The incidence of PVH/IVH has declined over the past several decades, but is still estimated to occur in approximately 20-25% of all preterm births. It is most prevalent in infants born at less than 32 weeks GA (21). PVH/IVH can be caused by hypoxic ischemic encephalopathy, respiratory distress, or circulatory problems, and is graded I, II, III, or IV based on localization of the bleeding, the amount of blood in the ventricles and degree of ventricular distension (grade I is mildest, while grade IV is most severe).

The risk of neurodevelopmental disability increases with increased grade of PVH/IVH (23). Over 90% of children with a grade IV intraventricular hemorrhage will be diagnosed with a neurodevelopmental disability—such as cerebral palsy—later in life (21).

It is important to note that many infants with neurological sequelae do not have identifiable brain injury, such as PVL or PVH/IVH, on cranial ultrasound (24). In a preterm baby, CNS maturation designed to take place in the womb instead occurs in an extrauterine environment and this process may not be trouble-free, even in the absence of brain injury. Thus, neurodevelopmental disability and delay seen in some preterm infants may be the result of abnormal brain development rather than an identified brain injury or lesion (25). Conversely, many preterm children with white matter lesions have no neurological abnormalities. This is probably due to the incredible plasticity and adaptability of the immature brain, which—as described in a 2007 review by Wyatt—may occur through “remodeling of existing white and grey matter regions, the refinement and selection of dendritic connections, the rerouting of white matter tracts to circumvent obstructions, and the development of alternative cortical processing strategies” (19). Although cranial ultrasound and MRI give the clinician important information on brain structures, they do not provide insight on neurological function. Thus, the results of neuroimaging must be combined with clinical neurological assessment to enhance clinical decision-making.

1.4 Neurodevelopmental outcomes in preterm infants

As the number of preterm births in the United States has risen, there have been significant advances in high-risk obstetric and neonatal care. As a result, through the 1990s, preterm infant mortality rates decreased and the limits of viability lowered to approximately 23-24 weeks gestational age (2, 4, 5, 7, 26). As survival rates improved, there was little change in the incidence of major medical complications—such as brain injury, necrotizing enterocolitis, chronic lung disease, and poor postnatal growth—predictive of future disability in preterm infants, and the number of preterm children with poor neurodevelopmental outcomes rose sharply (2, 3, 5, 27, 28). In the past decade, neonatal morbidities have begun to decrease, largely in part to improved medical care, and there has been a leveling off or perhaps even a slight decrease in the proportion of preterm children with significant neurodevelopmental impairment and cerebral palsy (6, 7). However, due to increasing preterm birth rates as well as improvements in detection of minor neurological dysfunctions such as learning disabilities, behavioral problems, and developmental coordination disorder, the number of preterm children with identified neurodevelopmental dysfunction continues to rise (3, 29-40).

Cerebral palsy (CP)—one of the most severe and disabling consequences of prematurity—results from neurological lesions, such as intraventricular hemorrhage or periventricular leukomalacia, suffered prior to, during, or soon after birth (20). Studies show that the prevalence of CP in preterm infants varies greatly with factors such as gestational age, birth weight, and medical complications. Depending upon

these factors, CP is diagnosed in anywhere from 7% to over 20% of preterm infants (3, 20, 27, 32, 37, 41-45). The developmental course after brain injury in preterm infants is highly variable and poorly understood. Transient neurological signs—such as dystonia—result in high rates of false-positive results on early neurological assessments (children who are neurologically normal test in the abnormal range). To avoid misdiagnosis, most children are not labeled as having CP or other major neurological disorders until they are nearly 2 years old. Children with mild forms of CP may not be diagnosed until even later in childhood (3). Although most preterm infants will not be diagnosed with CP, many demonstrate moderate neuromotor dysfunction or exhibit significant cognitive deficits by school age (3, 6, 35, 36, 41, 43, 45-47).

There has been growing interest in “high prevalence/low severity” neurological dysfunction—or “soft” signs—in children born prematurely. These more subtle dysfunctions occur in high rates in children who were preterm, and include learning disabilities, minor motor impairment such as developmental coordination disorder, subnormal IQ, and social-emotional and behavioral problems (3, 21). In fact, studies have shown that the majority of preterm infants have below-average fine and gross motor skills at school age, and they tend to score consistently lower on educational testing compared to their term counterparts (3, 30, 33, 34, 38, 39, 47-52). Because these minor dysfunctions often involve difficulties with higher-level skills, they are usually not recognized until school age, when cognitive and developmental demands are higher than in early childhood (21, 31). Of particular

interest to physical therapists is the recent increase in the incidence of developmental coordination disorder (DCD) among preterm children. Recent studies have shown that anywhere from 10-31% of extremely low birth weight or very preterm children have DCD, which is characterized by poor motor performance in the absence of global developmental delay or neurological signs associated with CP (50, 52). Like other low-severity dysfunctions, DCD is often not diagnosed until school age, if ever, and is often seen in otherwise “normal” preterm children (53). This begs the question of whether there may be early, transient neurological signs that could be assessed in preterm babies that would provide an early indication of this future mild motor dysfunction.

1.5 Predictors and risk factors for poor neurodevelopmental outcomes

Prediction of preterm infant outcomes remains one of the most challenging aspects of neonatal care. Although studies have identified some risk factors and predictors of poor developmental outcomes in groups of infants, individual outcomes are much more difficult to predict. This is in part because—in addition to perhaps more obvious predictors of poor neurodevelopmental outcomes such as abnormal cranial ultrasound or abnormal neurological exam—there are several non-neurological factors such as gestational age at birth, gender, and respiratory history that influence risk.

Neuroimaging and neurological examination results are two of the most studied risk factors for future neurodevelopmental delay. Multiple studies have

shown that infants with grades III or IV IVH or PVL on cranial ultrasound have a significantly greater risk of moderate neuromotor dysfunction or more severe deficits such as cerebral palsy (21, 23, 28, 46, 47, 54, 55). However, as described in Section 1.3, there are many infants with normal cranial ultrasound findings that go on to exhibit neurodevelopmental disabilities and delays and there are a large proportion of infants with abnormal cranial ultrasound findings that have little to no functional neurological deficits (19, 24, 25, 49). In 2005, Rademaker and colleagues reported that, in preterm infants, MRI showed more subtle white matter abnormalities and correlated more strongly with outcomes than cranial ultrasound. This suggested that cranial ultrasound, which is much more commonly used in the NICU than MRI, may not be sensitive enough to detect brain lesions in all preterm babies (56). But it is also possible that developmental outcomes have as much to do with brain development and adaptation as with specific brain lesions. Thus, although results of neuroimaging may provide one model of prediction, those results should not be viewed as diagnostic (43). Recent studies have suggested that clinical neurological examination may be as significant as neuroimaging in predicting developmental outcomes and that the two should be combined to optimize prediction of poor neurodevelopmental outcomes (57, 58).

There are a number of non-neurological factors that have also been found to influence risk. Infants born at younger gestational ages (GA) have worse neurodevelopmental outcomes than older preterm or term infants (3, 21, 43, 59). However, it is difficult to determine whether the increase in risk is truly due to

decreased birth GA or due to medical complications frequently associated with younger preterm babies. Bronchopulmonary dysplasia—defined as oxygen dependence at 36 weeks PMA—is one such complication that has been highly associated with poor outcomes in preterm infants. Infants with bronchopulmonary dysplasia have demonstrated increased rates of developmental delay through 18 months adjusted age in a variety of studies (28, 60, 61). This may be due to the increased risk of hypoxia in infants with respiratory distress or because respiratory insufficiency may be an indicator of overall sickness, which is more likely to result in poor outcomes. Other potential risk factors for neurodevelopmental delay—according to the literature—include retinopathy of prematurity, male sex, low birth weight, and low Apgar scores (24, 28, 31, 33, 37, 38, 46, 50, 62, 63). Much of the evidence contends that the best prediction models do not look at risk factors in isolation, but rather consider a combination of some or all neurological and non-neurological risk factors (21, 28, 54, 62).

1.6 Effects of early intervention

While in the NICU or upon NICU discharge, preterm infants are frequently referred for early intervention services such as physical, occupational, or speech therapy. A 2004 study by Clements and colleagues revealed that, among preterm children in the state of Massachusetts, the cost of early intervention services averaged \$5393 per infant born at 24-31 weeks gestational age (GA) and \$1578 per infant born at 32-36 weeks GA, compared to just \$725 per term infant (64). Despite the large

amounts of dollars that states like Massachusetts are spending on habilitation of preterm children, there is conflicting evidence whether early intervention services are effective in improving outcomes in this population (8-10).

A recent meta-analysis by Spittle and colleagues reviewed 16 randomized controlled trials of early intervention that began in the first year of life for preterm infants, and found that early intervention improved cognitive outcomes in infancy and preschool but not at school age. Early intervention had no effect on motor outcomes at any age (11). This seems to be discouraging news for early interventionists working to improve motor outcomes in preterm infants. However, it is important to note that many of these studies recruited only low-risk infants or included very broad samples of preterm babies (i.e. very low birth weight or birth prior to 33 weeks gestation age) without regard for risk and, therefore, possible need for these services. Two studies have looked at the efficacy of early motor interventions specifically provided to high-risk infants, and both found positive effects.

In 1994 Girolami and Campbell (12) recruited infants born at ≤ 35 weeks GA who also had at least one of the following complications: 5-minute APGAR ≤ 5 , intraventricular hemorrhage, seizures, respiratory distress syndrome, respiratory arrest, birth weight < 1000 grams, central nervous system depression or irritability, asphyxia, mechanical ventilation, or thermal instability. These babies were then tested using the Neonatal Behavioral Assessment Scale (NBAS) and were only included in the study if they demonstrated at least 3 abnormal or asymmetrical reflexes. Identified at-risk infants were randomized into control and treatment

groups. Infants in the control group received no physical therapy intervention in the NICU. Infants in the treatment group received neurodevelopmental treatment (NDT)-based intervention twice per day for 7-17 days. A full-term comparison group received no intervention. At the end of the study, the preterm treatment group performed significantly better on the supplemental motor test (SMT), a test designed for the study to assess functional postural control. This suggests that high-risk preterm infants may benefit from physical therapy in the NICU.

Lekskulchai and Cole (13) published a similar study in 2001, but looked at a larger group of preterm infants and measured outcomes over a longer time period. For their study, they tested all participants using the Test of Infant Motor Performance (TIMP). Infants with scores less than 67 were designated as “at-risk” and randomly assigned to a treatment group or control group. The control group did not receive intervention and the treatment group received developmental physical therapy intervention from term to 4 months adjusted age. Infants with scores of greater than 67 were considered to be not at-risk and were included in a comparison group that received no intervention. At term, both the intervention and control groups had worse TIMP scores than the comparison group. However, by four months adjusted age, the intervention group scored significantly higher on the TIMP than both the preterm control group and the comparison group. The results of this study strengthen the argument that early intervention physical therapy services may indeed be beneficial when targeted at high-risk preterm babies. Reliable and valid clinical

neurological assessment tools would assist the clinician in identifying those high-risk babies that would be most likely to benefit from early intervention services.

1.7 The role of clinical neurological assessment in identification of at-risk infants

The neonatal neurological examination is designed to measure the function and maturity of the central nervous system. Most standardized neonatal neurological examinations include items that assess posture, movement, active and passive muscle tone, primitive reflexes, and/or responsiveness to sensory stimuli. Although physicians frequently assess these items in a non-standardized fashion as part of routine neonatal care, use of a standardized assessment tool has many benefits. Standardized neurological assessments may aid the clinician in establishing a baseline for neurological function and, when administered serially, may reveal infants with transient, resolving neurological abnormalities versus infants with persistent abnormalities. Furthermore, standardized neonatal assessment tools may—if psychometrically sound—be used as an outcome measure for research purposes (65). Recent studies have shown that standardized neonatal neurological assessments in the NICU are effective in predicting neurodevelopmental outcomes, particularly when used in combination with neuroimaging results (57, 58). Predictive standardized assessment tools can aid the clinician in clinical decision-making when making referrals for early intervention services and planning for NICU discharge and follow-up.

Currently, a number of clinical neurological assessments for preterm infants are available for use by neonatal health care providers and researchers. However, there is no agreed-upon “gold standard” for preterm infant neurological assessment. Psychometric soundness is not well established for many assessments for preterm infants. This is largely because—in preterm babies—autonomic instability leads to disorganized state- and self-regulation, resulting in inconsistent, sometimes abnormal motor control and neurobehavior (66). Often, even when reliability and validity of standardized assessments seem to be clinically significant, wide variations both between and within babies make statistical significance nearly impossible to achieve. This is particularly true for very sick preterm babies and/or those at younger post-menstrual ages, when the CNS is stressed and immature and self-regulation is particularly poor. As sickness resolves, babies become more medically stable, and as they approach term-adjusted age, functional neurological status is less confounded with physiological status (67) and results of standardized testing may become more consistent. As a result, researchers studying psychometrics of standardized neurological assessment tools frequently solve the problem of infant variability and relatively low statistical relationships by testing only older preterm babies and/or low-risk, “healthy” preterm infants with few or no major medical complications or risk factors. However, this practice results in tests that are validated on only a fraction of infants who may be admitted to the NICU and speaks little to the clinical significance of such assessments. When sicker or younger preterm infants are studied, lower

correlations may be used to represent meaningful, clinically significant—although not necessarily statistically significant—relationships.

While tools that do have established psychometric properties are useful for research purposes, they are often not clinically useful as they may take a great deal of time to administer and/or score and they often require extensive training. For example, the Assessment of Preterm Infant Behavior (APIB) is a neurobehavioral assessment for preterm infants who can be handled on room air (68). The authors of the APIB report that it has a very sound theoretical and scientific basis and that interrater reliability is easily established in trained examiners (69). Studies on the APIB have shown acceptable sensitivity and construct validity. However, it requires 8-10 months of training for the advanced clinician and takes up to 60 minutes to administer, 30 minutes to score, and 3 hours to write a summary report (68, 69).

The Neurobehavioral Assessment of the Preterm Infant (NAPI) requires that examiners view a training videotape, and it takes only about 30 minutes to administer (68). The NAPI has well-documented reliability and validity. However, it can be administered only on infants at 32 weeks PMA or later (70-72). This limits its ability to be used for serial assessment of very preterm infants born before 32 weeks GA.

The Test of Infant Motor Performance (TIMP) can also be administered beginning at 32 weeks PMA, so its use is limited in younger preterm babies. It requires 36 minutes to administer and score (73). The TIMP is a relatively new assessment, and its psychometric properties continue to be studied. It appears to have good test-retest reliability and it is responsive to change over time. However, there

are questions about the construct validity of the TIMP as there is conflicting research on whether it discriminates among infants at varying risk for poor neurological outcomes (73-75).

The Neonatal Intensive Care Unit Network Neurobehavioral Scale (NNNS) is another relatively new neurological assessment for preterm infants. It was originally designed for a study on prenatal drug exposure (76). The NNNS requires certification through a special training program, and the test can be completed in less than 30 minutes on infants at 30 weeks PMA or later (76, 77). Test-retest reliability has been established in preterm infants. However, validity has only been established in term infants (76).

1.8 The Premie-Neuro

The Premie-Neuro is a standardized clinical neurological assessment for preterm infants 23-37 weeks PMA. It consists of 3 subtests, each containing 8 items. The Neurological subtest contains items that assess primitive reflexes such as plantar or palmar grasp, muscle tone (i.e. popliteal angle and scarf sign), and movement type. The Movement subtest tracks the rate per minute of behavioral signs of stress and avoidance behavior such as yawning, color changes, and limb movements. Finally, the Responsiveness subtest assesses active tone (head and trunk control), alertness, and responsiveness. Because the Responsiveness subtest involves changing the position of the infant, it is not recommended for infants who are less than 28 weeks

PMA or who are on a ventilator. Thus, for these more immature and sicker babies, only the Neurological and Movement subtests are administered (78).

The Premie-Neuro was first described in 2005 by Daily and Ellison, who identified few available methods for assessing brain function in preterm infants, particularly in those born at less than 28 weeks gestational age. They constructed the tool to be “acceptable to the clinician in terms of the stress it places on the sick or convalescing newborn, the time required for assessment, the clarity of items, and the simplicity of scoring” (14). In their only published paper on the tool, Daily and Ellison described the construction of the 24-item assessment from 75 original items. Internal consistency of the tool ranged from 0.73-0.82 (14). A pilot study of the Premie-Neuro reported significant predictive validity of the tool at term and 6-8 months adjusted age (79).

The Premie-Neuro is quick to administer, requiring less than 10 minutes. It does not require extensive examiner training and can be done with minimal disturbance of the infant, without removing him from his crib or isolette and without interfering with electronic monitoring or mechanical ventilation. At this time, it is the only standardized neonatal assessment that can be safely administered on infants as young as 23 weeks PMA (14, 78).

1.9 Significance of presented work

Repeated neurological assessments of preterm infants are critical in identifying neurological abnormalities and determining if those abnormalities persist

or resolve (16, 68). Unfortunately, many currently available assessments are not optimal for repeated examination because of the time required to administer and/or the limited age range in which the test may be performed. The Premie-Neuro is designed to overcome these problems, making it a promising tool for serial neurological assessment in the NICU. Administration of the Premie-Neuro does not require extensive training. Healthcare providers with experience handling and observing preterm infants may be trained by simply reviewing the administration manual and observing the testing procedure. The Premie-Neuro can be administered in less than 10 minutes, requiring much less time than many comparable assessment tools and requiring much less handling and potential stress to the infant. The Premie-Neuro can be safely completed on an infant as young as 23 weeks PMA, even if that infant is on a ventilator and/or requires electronic monitoring (14). This means that the Premie-Neuro can be administered on even the youngest and sickest infants, and can continue to be used as the infant's health improves and the infant reaches term adjusted age (37 weeks PMA). Despite these unique features of the Premie-Neuro, it is not yet widely used because—until now—the psychometric properties of the assessment have not been established. In the present study, we will present evidence that Premie-Neuro raw score is reliable and valid. These findings may change the way neonatal neurological assessment is conducted in the NICU.

The present study may change the way early intervention services are provided in the NICU and following NICU discharge. Currently, there are few available methods to identify at-risk infants in the NICU, so developmental

interventions are often postponed until evidence of a delay or disability is undeniable. However, there is consensus that one of the most rapid periods of brain growth and development is the last 3 months preterm through the first several months post-term (8, 16, 19). Although there is conflicting evidence whether early intervention services are effective in improving outcomes in all babies, it is difficult to analyze these studies due to the large variability in the subject populations studied, methods of intervention used, and outcome measures. There are several well-conducted studies that show clear benefits of developmental care (such as NIDCAP) and task-specific training for low-risk preterm babies, as well as a positive effect of physical therapy interventions selectively administered to high-risk babies (8, 12, 13, 80-82). Perhaps the evidence will strengthen as we begin asking the right questions or using appropriate outcome measures, but it is clear that interventions provided for high-risk babies in the first months of life can impact functional outcomes. The presented work may impact utilization of early intervention services for preterm babies in two ways. First, in Chapter 4, we will suggest modifications of Premie-Neuro cut-points used to classify infants as normal or not normal. This classification system will allow earlier identification of at-risk infants—those who repeatedly score in the not normal range—so that those babies may be evaluated further and/or targeted for intervention services at earlier ages. This will also reassure parents and clinicians of infants who are repeatedly classified as normal, and may prevent over-utilization of services for infants who do not demonstrate any functional neurological abnormalities. Second, this work will allow clinicians and researchers the ability to quantify neurological

development at younger ages than ever before so that the effectiveness of interventions provided as early as 23 weeks PMA in the NICU may be monitored or formally studied.

1.10 Specific aims and statement of hypotheses

The purpose of this presented work was to determine the reliability and validity of the Premie-Neuro for preterm infants in the neonatal intensive care unit (NICU).

Specific Aim 1: Determine the **reliability** of the Premie-Neuro in preterm infants in the NICU.

Reliability is a measure of the repeatability of a test when given under identical conditions (65). High reliability does not necessarily indicate a “good” test, but does indicate that there is minimal error resulting from variability between or within raters (83). Preterm infants demonstrate large variations in neurological function while in the NICU, which often proves challenging when attempting to identify high-risk infants and predict neuromotor outcomes (3). It is important that a standardized clinical neurological assessment have acceptable interrater and test-retest reliability to ensure that variation in scores is due to true variation in the infant’s neurological status and not due to error variance of the examiner or the instrument. Thus, we hypothesized that interrater reliability of Premie-Neuro raw score would be high ($ICC \geq 0.80$) and that one-day test-retest reliability of Premie-

Premie-Neuro raw score would be high ($ICC \geq 0.73$). We further hypothesized that Premie-Neuro classifications (normal, questionable, and abnormal) would demonstrate at least moderate agreement ($Kappa > 0.40$) over a one-day period, and that Premie-Neuro classifications would be stable over at least two consecutive weeks, with infants remaining in one classification for at least 60% of all assessments.

Specific Aim 2: Establish **construct validity** and **responsiveness** of the Premie-Neuro in preterm infants in the NICU.

Validity refers to the extent that an instrument measures what it is intended to measure (65). Responsiveness describes a test's ability to detect clinically important change (68, 75, 84). We believed that the Premie-Neuro would detect differences between two groups of infants known to be at high and low risk for neurological insult and that it would be responsive to different rates of neuromaturational change in infants in these groups. Specifically, we hypothesized that Premie-Neuro raw scores would be significantly lower (worse) for high-risk infants, and that low-risk infants' scores would remain unchanged while high-risk infants' scores decreased over time. We further hypothesized that a larger proportion of high-risk infants would be classified as neurologically questionable or abnormal versus low-risk infants. Significance was set at $p < 0.05$.

Construct validity of the Premie-Neuro was also analyzed by determining if characteristics that are associated with poor neurological outcomes are also associated with Premie-Neuro scores. We hypothesized that low GA at birth, presence of CNS

injury and diagnosis of bronchopulmonary dysplasia (BPD)—characteristics that are associated with poor neurodevelopmental outcomes (23, 24, 28, 37, 43, 54, 55, 59-61)—would be significant ($p < 0.05$) predictors of Premie-Neuro neurological scores and classifications.

Specific Aim 3: Estimate the **predictive validity** of the Premie-Neuro in preterm infants in the NICU.

Predictive validity is the ability of performance measured by one instrument to predict future performance on another instrument that measures the same construct (65, 83). This type of validity is particularly relevant to clinicians in the NICU as it allows them to make a neurological prognosis, identify appropriate interventions, and plan for follow-up care. In this study, we hypothesized that Premie-Neuro scores obtained before 38 weeks PMA would have a positive and fair (>0.25) correlation with NeoNeuro scores at term and Infanib and AIMS scores at 3 months adjusted age. We further hypothesized that there would be a significant ($p < 0.05$) relationship between Premie-Neuro classifications obtained in the NICU and NeoNeuro classifications at term and Infanib and AIMS classifications at 3 months adjusted age.

1.11 Summary

The last 3 months of gestation may represent a critical period in brain development as it is a time of rapid synaptogenesis and myelination (16-19). Unfortunately for the preterm infant, this time is spent in the NICU rather than the

womb. It is true that the preterm infant brain is more susceptible to lesions such as intraventricular hemorrhage and periventricular leukomalacia, but—even in the absence of these lesions—preterm infants exhibit neurodevelopmental delay and disability resulting from suboptimal CNS development (19-21, 25, 27). Although neuroimaging provides invaluable information on brain structure, this information must be combined with clinical data, such as results of standardized clinical neurological assessments, to determine brain function, identify at-risk infants, and improve prediction of those babies who will have poor neuromotor outcomes (57, 58). The findings of the presented work may allow clinicians to reliably assess preterm infants and combine those findings with neuroimaging results and clinical data to identify at-risk infants, make decisions about referring for specialized early intervention services, and predict neuromotor outcomes at younger ages than ever before.

Chapter 2

The Premie-Neuro:

A reliable and valid clinical neurological test for preterm infants in the NICU

2.1 Abstract

There are few methods available for clinicians to assess the functional neurological status of the newborn. As more preterm infants continue to live, there is a growing need for reliable and valid clinical neurological assessment tools to assist the clinician in targeting preterm infants at risk for developmental delay. The purpose of the present study was to determine the reliability and validity of the Premie-Neuro, a new clinical neurological test for preterm infants. In this study, infants were assessed twice—24-72 hours apart—using the Premie-Neuro. One of these assessments per infant was observed and scored by a second examiner. Infants continued to be administered the Premie-Neuro weekly through 37 weeks post-menstrual age or discharge from the NICU. At discharge, infants' medical charts were reviewed to determine medical risk factors for neurodevelopmental delay. Infants were categorized as “high” (high risk, HR) or “low” risk (low risk, LR) for neurodevelopmental delay using the Neurobehavioral Risk Scale (85). The Premie-Neuro had fair to moderate interrater and one-day test-retest reliability. Premie-Neuro classifications were stable across weeks. Premie-Neuro scores were higher in the LR group. Scores tended to worsen over time in the HR group but not in the LR group. Infants who were not oxygen-dependent at 36 weeks post-menstrual age were more likely to be classified as neurologically “normal” using the Premie-Neuro. Oxygen dependence at 36 weeks post-menstrual age and CNS abnormalities (grades III and IV intraventricular hemorrhage or periventricular leukomalacia) were significant predictors of Premie-Neuro score. These data suggest that the Premie-

Neuro raw scores have acceptable reliability and validity for clinical use with preterm infants. Premie-Neuro classifications should be interpreted with care. The Premie-Neuro may be used as an adjunct to clinical assessment to identify preterm infants at-risk for developmental delay.

2.2 Introduction

Over 12% of infants in the United States are born too soon. The rate of preterm births—defined as birth prior to 37 completed weeks gestational age (GA)—has increased significantly over the last several decades (1). Advances in modern medicine, including the introduction of antenatal steroids and postnatal surfactant, led to improved survival rates among preterm infants through the 1980s and 1990s. But increased survival was accompanied by increased neonatal morbidities, and the number of preterm infants with resultant neurological sequelae—such as cerebral palsy—rose through the mid- to late-1990s (2-5, 36). Over the last decade, as survival rates have leveled off and medical care has improved, the proportion of preterm children with significant neurodevelopmental disability or delay has not increased (6, 7). However, with rising rates of prematurity and improved detection of more subtle neurological disabilities, there continues to be large numbers of preterm children who may benefit from specialized care in the neonatal intensive care unit (NICU) and throughout childhood. A recent review by Blauw-Hospers and Hadders-Algra found that developmental care, such as the Newborn Individualized Developmental Care and Assessment Program (NIDCAP), improved outcomes among the general population of preterm children (8). However, more specialized early interventions—such as physical therapy—do not seem to affect outcomes unless they are targeted specifically at preterm infants who have already been identified as at-risk for neurodevelopmental delay or disability (9, 10, 12, 13). This suggests it may not be appropriate to refer a preterm infant for early intervention services simply

because he is preterm, and speaks to the growing need for clinical neurological assessment tools that would reliably identify high-risk infants, assisting the clinician in making appropriate referrals for early intervention and improving efficiency and avoiding over-utilization of these services.

Although there are a variety of assessment tools available for use in the preterm infant population, there is no agreed-upon “gold standard” for neonatal neurological evaluation. Many tests are quite lengthy to administer, potentially placing prolonged stress on the infant. Furthermore, many clinical neurological assessment tools require extensive examiner training and few are appropriate for infants younger than 30-32 weeks post-menstrual age (PMA). It is crucial that health care providers have access to a clinical neurological assessment tool to identify high-risk preterm infants in the neonatal intensive care unit (NICU) so that early intervention services can be initiated, if necessary, during the critical period of rapid brain maturation that occurs in the first months of life. It is clinically important that such a tool requires minimal infant handling, contains items appropriate for all preterm infants in the NICU—even the tiniest infants who may be as young as 23 weeks PMA—and is quick and relatively easy to administer. The Premie-Neuro, a standardized neurobehavioral assessment tool developed by Donna Daily, MD, has the potential to fill this role (14).

The Premie-Neuro consists of 3 subscales: Neurologic, Movement, and Responsiveness. Each subscale consists of eight items for a total of 24 items on the full Premie-Neuro (Table 2.1). The Neurologic subscale includes reflexes and

measures of passive tone such as palmar and plantar grasp, arm recoil, and popliteal angle. The Movement subscale requires that the examiner track and calculate the rate (per minute) of spontaneous limb movements, color changes, startles, and other physiologic and behavioral stress signs during the examination. The Responsiveness subscale includes measures of active tone, head and neck control, alertness, and responsiveness when the infant is placed in a variety of positions such as supported sitting and ventral suspension. Infants who are 28 weeks PMA or younger or who are on a ventilator are only administered the Neurologic and Movement subscales of the Premie-Neuro, a total of 16 items. Infants older than 28 weeks PMA and not on a ventilator are administered all 24 items in the 3 subscales. Each item is assigned a score of one, three, or five based on the infant's gestational age and performance on the item. Item scores are tallied and the infant is classified as neurologically normal, questionable, or abnormal based on raw score cut-points recommended by the Premie-Neuro authors in *The Premie-Neuro Examination Instruction Manual* (78).

In 2005, Daily and Ellison described the construction of the Premie-Neuro and reported its internal consistency to range from 0.73-0.82 (14). A pilot study of the Premie-Neuro reported acceptable predictive validity at 6-8 months adjusted age (79). However—although the tool is in limited use—complete psychometric properties of this assessment, including reliability and validity, have not yet been established. Thus, the purpose of this study was to establish the interrater and test-retest reliability as well as construct validity of the Premie-Neuro.

2.3 Methods

Participants

Thirty-four preterm infants (mean gestational age at birth 29 ± 3.67 SD weeks, BW 1343.18 ± 696.25 SD grams) were recruited from the NICU (Table 2.2). In accordance with the Human Subjects Committee, each participant's parent or guardian signed an institutionally approved parental consent form. Infants were included in the study if they were born prior to 37 weeks gestation and were no more than 37 weeks PMA at the time of the first Premie-Neuro assessment. Infants with major medical complications were included in the study if their physician agreed that they were medically stable enough to undergo clinical neurological assessment. Infants born at or after 37 weeks completed GA or with a known genetic disorder or major congenital anomaly were excluded from the study. In order to demonstrate significance at a power of 0.80, a minimum of 10 infants was needed to test interrater reliability, and approximately 30 infants were needed for all other analyses.

Procedure

Premie-Neuro testing began as soon as possible after obtaining parental consent. Prior to each test, the infant's physician or nurse was consulted to ensure that the infant was sufficiently medically stable to undergo hands-on assessment that day. All Premie-Neuro testing was conducted at bedside by a physical therapist. Each infant remained in his or her crib or isolette during testing. In order to minimize the potentially confounding effect of alert state, Premie-Neuro assessments were

conducted within the hour prior the each participant's scheduled feeding time whenever possible. Signs of physiological stress were monitored during testing. Testing was discontinued and/or canceled and rescheduled if the infant showed significant signs of stress—such as elevated heart rate or low oxygen saturation—during the testing or was acutely ill the day of testing.

During the first week of the study, the Premie-Neuro was administered two times by the same examiner (Examiner 1, E1) 24-72 hours apart. During the second assessment, a second examiner (Examiner 2, E2) observed the hands-on assessment and independently scored the Premie-Neuro. E1 continued to administer the Premie-Neuro approximately weekly (once every 5-9 days) through 37 weeks PMA or discharge, whichever occurred first.

At the time of enrollment in the study, each infant's birth weight, sex, and gestational age at birth were obtained by reviewing the infant's medical chart. At discharge, the infant's medical chart was reviewed to obtain the following information: blood pH history, seizure activity, evidence of central nervous system injury, respiratory history (including need for mechanical ventilation), and history of hypoglycemia. These data were used to calculate a Neurobiologic Risk Score (NBRS) (85, 86) for each infant. Infants with a discharge NBRS of <5 were considered "low risk" (Low Risk Group, LR) and infants with a discharge NBRS of ≥ 5 were considered "intermediate/high risk" (High Risk Group, HR) for neurological injury and poor neurodevelopmental outcomes (85, 87). Descriptive information for HR and LR groups is included in Table 2.2.

Data analysis

Intraclass correlation coefficients (ICC) were calculated for raw scores obtained by E1 and E2 for the same assessment to estimate interrater reliability and for scores obtained by E1 on consecutive days to estimate short-term test-retest reliability of the Premie-Neuro. Reliability of Premie-Neuro classifications (normal, questionable, or abnormal) was calculated using Kappa agreements. Early sickness and autonomic instability in the young preterm infant can manifest as disorganized sensory and motor function (66). Because of this, there is significant variability in the neurobehavioral status of preterm babies—both within and between infants—particularly at young gestational ages (67). Thus, in preterm infants, lower correlations—in the fair to moderate range—may be used to signify clinically important relationships. For this study, reliability calculations were interpreted cautiously using Landis and Koch criteria: <0.20 slight, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, and 0.81-1.00 almost perfect (88). Construct validity and responsiveness of the Premie-Neuro were assessed using a two factor (Group by Post-Menstrual Age) repeated measures ANOVA with Premie-Neuro raw score as the dependent variable. Fisher's exact test was used to determine if there was a significantly larger proportion of LR infants classified as normal on the Premie-Neuro versus HR infants at both the first and discharge Premie-Neuro assessments. Logistic and linear regression analyses with stepwise elimination were used to determine if GA at birth, significant CNS injury (grades III or IV intraventricular hemorrhage, periventricular leukomalacia, residual ventriculomegaly, and intraparenchymal

periventricular echodensities) or oxygen dependence at 36 weeks were predictive of Premie-Neuro classifications and raw scores, respectively, at first and discharge Premie-Neuro assessments (2, 23, 28, 54, 55, 60, 61).

In accordance with the Premie-Neuro instruction manual (78), all infants were administered the Neurologic and Movement subscales of the test. Infants older than 28 weeks PMA and who were not on a ventilator at the time of testing were also administered the Responsiveness Subscale. Thus, selected analyses were conducted for two subtests for all infants and for all three subtests only for infants who were administered the full Premie-Neuro. Analyses were performed with SPSS 16.0*. Significance was set at $p < 0.05$.

2.4 Results

Interrater reliability of the Premie-Neuro

Interrater reliability data were collected from a subset of 15 infants. The average PMA at the time of interrater reliability assessment was 33.47 weeks (± 1.73 SD weeks). All fifteen infants were administered the Neurologic and Movement subtests. Mean total raw score for these two subtests was 67.20 (± 3.99 SD points) for the hands-on examiner (E1) and 66.53 (± 4.17 SD points) recorded by the observer (E2). Interrater reliability was moderate for total raw score for two subtests (ICC=0.556, Table 2.3). One infant was on a ventilator at the time of interrater reliability testing and was not administered the Responsiveness subtest. Thus,

xxxvi—

* SPSS, 233 S. Wacker Drive, Chicago, Illinois, 60606

interrater reliability data for all 3 subtests were obtained from 14 infants. Mean total raw score for all three subtests was 99.14 (± 4.62 SD points) for the examiner and 99.29 (± 4.34 SD points) recorded by the observer. Interrater reliability was fair for total raw score for all three subtests (ICC=0.391, table 2.3). For Premie-Neuro classifications obtained for 15 infants by different examiners (E1 and E2) for the Neurologic and Movement subtests, $\kappa=0.400$ ($p=0.121$). For classifications for the fourteen infants administered all three subtests, $\kappa=0.143$ ($p=0.593$, Table 2.3).

Test-retest reliability of the Premie-Neuro

One-day test-retest reliability was collected from 30 infants (mean PMA at assessment 33.07 ± 2.26 SD weeks). Four infants were not included in this analysis because of NICU discharge prior to administration of the second Premie-Neuro test. All 30 infants were administered the Neurologic and Movement subtests. Mean total raw score for these two subtests was 68.40 (± 4.80 SD points) on day 1 and 66.20 (± 5.57 SD points) on day 2. One-day test-retest reliability for two subtests was moderate (ICC=0.493). Four infants were on a ventilator at the time of short-term test-retest reliability testing, so 26 infants were administered all 3 subtests of the Premie-Neuro. Mean total raw score for infants who were administered all three subtests was 100.54 (± 6.59 SD points) obtained on day 1 and 99.85 (± 5.85 SD points) obtained on day 2. Reliability of total raw score for all three subtests was moderate (ICC=0.592). For the 30 infants who were administered the Neurologic and Movement subtests of the Premie-Neuro on consecutive days by E1, $\kappa=0.429$

($p=0.018$). For the 26 infants who were administered all 3 subtests, $\kappa=0.133$

($p=0.484$). Test-retest reliability data are summarized in Table 2.3.

Stability of Premie-Neuro classifications

Eight infants were discharged from the NICU before the second week of testing. For the 26 infants who were administered the Premie-Neuro on at least two consecutive weeks, the total number of Premie-Neuro assessments ranged from 2-10 (mean= 4.69 ± 2.24 SD assessments). No infants tested in the abnormal classification throughout the course of the study. Sixteen infants (61.54%) were classified as questionable for the majority of their assessments. Eight infants (30.77%) were classified as normal for the majority of their assessments and 2 (7.69%) were classified as normal and questionable for an equal number of assessments. Overall, infants tested in their most frequent classification an average of 79.52% of all assessments (Table 2.4).

Comparison between high-risk and low-risk infants

All 34 subjects were administered at least one Premie-Neuro assessment in the NICU. Using the NBRS, eleven infants were categorized as high risk (high risk group, HR) and the remaining twenty-three infants were categorized as low risk (low risk group, LR). As expected, scores were lower in the HR group versus the LR group at both the first and the discharge Premie-Neuro assessment (Table 2.5). As shown in Table 2.6, there was a significant main effect of both post-menstrual age and risk

group on Premie-Neuro raw scores, and a significant interaction between risk group and post-menstrual age on Premie-Neuro raw scores. Although scores for both groups declined over time, scores for infants in the HR group tended to decline at a much faster rate than infants in the LR group. While scores for infants in the LR group tended to level off at approximately 32 weeks PMA and remained stable through term adjusted age, scores for infants in the HR group continued to fall as the infants approached term adjusted age (Figures 2.1 and 2.2). There was not a significantly higher proportion of LR infants classified as normal versus HR infants at either the first ($p=0.252$) or discharge ($p=0.053$) Premie-Neuro assessment.

Relationship between risk factors and Premie-Neuro raw scores and classifications

Oxygen-dependence at 36 weeks PMA and evidence of central nervous system injury were significant predictors of discharge Premie-Neuro raw scores ($p=0.013$ and 0.024 , respectively). Only oxygen-dependence at 36 weeks was predictive of the discharge Premie-Neuro classifications ($p=0.038$). Gestational age at birth was not predictive of discharge Premie-Neuro scores or classifications and central nervous system injury was not predictive of discharge Premie-Neuro classifications (Table 2.7).

2.5 Discussion

Although the majority of infants born preterm do well, they are at much greater risk for developmental disability and delay than their term counterparts (2, 3,

21, 31, 33, 34, 38, 41-44, 46, 50-52). The risk of poor outcomes increases with neonatal medical complications such as bronchopulmonary dysplasia, central nervous system lesions, retinopathy of prematurity, and poor postnatal growth (21, 23, 24, 28, 32, 44, 46, 47, 54, 55, 60, 61). However, although these factors may predict outcomes in groups of infants, predicting individual outcomes remains difficult. There is mounting evidence that standardized developmental testing combined with informal clinical assessments and medical tests, such as MRI or cranial ultrasound, may be effective at identifying at-risk infants (57, 58). However, there are few reliable and valid tools available for the clinician to assess the functional neurological status of the newborn in the NICU, particularly at less than 32 weeks PMA. Our data show that the Premie-Neuro is a reliable and valid assessment tool for preterm infants, and may be used in combination with clinical examinations, general movement assessments, and/or neuroimaging to identify preterm infants most at-risk for developmental delay or disability.

Reliability values reported in this study for Premie-Neuro raw scores are slightly lower than those reported for many other assessment tools for preterm and newborn infants (65, 71, 72, 76, 89-92). However—in our study—infants were tested as young as 29 weeks PMA and few other instruments have been validated on such a young group of infants. Because preterm infants, particularly those at young post-menstrual ages, are extremely variable in terms of their developmental trajectory, high reliability is difficult to achieve in this population. In order to assess the true clinical reliability of the Premie-Neuro for the present study, the examiner and

observer did not participate in any formal training prior to beginning the study. Each simply read the instruction manual and watched the training video, as recommended by the authors of the Premie-Neuro. After the completion of the study, both the examiner and the observer reported that the instructions in the administration manual were often unclear or lacking sufficient detail. A more structured approach to training and a more detailed administration manual may enhance the interrater reliability of the Premie-Neuro. It is worth noting that our data for both test-retest and interrater reliability showed that mean total score varied little by day or between observer and examiner, respectively. In some cases, mean total raw score differed by less than one point. This shows that reliability of the Premie-Neuro may be more clinically significant than is reflected statistically. Although interrater or test-retest reliability of Premie-Neuro raw score is not substantial or near perfect according to Landis and Koch criteria (88), it is fair to moderate and therefore acceptable for clinical use in this highly-variable preterm infant population.

Interrater and one-day test-retest reliabilities of Premie-Neuro classifications are not as strong as the reliabilities of Premie-Neuro raw score. This is particularly true for classifications based on all 3 subtests versus classifications based upon only 2 subtests. Based on our results, we would suggest that the clinician use caution when interpreting Premie-Neuro classifications obtained over a short period of time (1-3 days). It may be more appropriate to interpret classifications over a longer period of time (one week or longer), as our data showed good, clinically significant stability of Premie-Neuro classifications over consecutive weeks.

To assess the construct validity of the Premie-Neuro, two methods were used. First, we assessed whether the Premie-Neuro could discriminate between two groups of infants known to differ in terms of risk for developmental disability and/or delay. As expected, HR infants had worse Premie-Neuro scores than LR infants, and the difference in Premie-Neuro raw scores between groups increased as infants approached term adjusted age. This is consistent with previous research that showed that, using the Test of Infant Motor Performance (TIMP), low-risk preterm infants tended to catch up with their full-term counterparts as they approached term and post-term adjusted ages, while high-risk preterm infants fell further behind over time (74, 92). Although the Premie-Neuro was excellent in discriminating between HR and LR infants in terms of raw score, it did not discriminate between HR and LR infants in terms of classifications.

Construct validity of the Premie-Neuro was also assessed by determining whether oxygen-dependence at 36 weeks PMA, central nervous system injury, and GA age at birth—all factors that have been shown to be predictive of later neurodevelopmental disability and/or delay—were predictive of Premie-Neuro scores and classifications. CNS injury and oxygen-dependence at 36 weeks PMA were predictors of Premie-Neuro total raw scores and only oxygen-dependence at 36 weeks PMA predicted Premie-Neuro classifications. This is consistent with literature showing that infants with abnormal cranial ultrasound findings and/or a significant history of respiratory illness and insufficiency tend to have worse long-term outcomes than healthy preterm babies (23, 28, 46, 47, 52, 54, 55, 60, 61). We were surprised to

find that, in the present study, central nervous system injury was not predictive of Premie-Neuro classifications. However, only six babies in the study had documented CNS injury, and the lack of significance may be due to the small sample size of this group. Gestational age at birth was not an independent predictor of Premie-Neuro scores or classifications, which is in agreement with other studies that have shown that—although children born at younger GA are at greater risk for poor neurodevelopmental outcomes—those outcomes are not as strongly tied to immaturity itself as they are to poor postnatal growth and medical complications that are often associated with extreme prematurity (2, 24, 32, 62).

Although Premie-Neuro raw scores were reliable and valid to a level that confirmed clinical relevance, reliability and validity of Premie-Neuro classifications were inconsistent and generally poor. We believe there are two possible explanations for this discrepancy. First, this study included a relatively small sample of preterm infants in the HR group and with CNS injury. It is possible that, due to these small sample sizes, our nonparametric statistical tests did not have sufficient power to detect differences. Another explanation—and one we believe is more likely—is that the currently recommended raw score cut-points for normal, questionable, and abnormal Premie-Neuro classifications may need to be modified. In the present study, a total of 118 full Premie-Neuro tests were administered on 34 infants. Overall mean Premie-Neuro score for these tests was 95.8. The mean Premie-Neuro total raw score for LR infants was 100.1, while the mean Premie-Neuro score for HR infants was 90. The current cut-points for Premie-Neuro classifications are abnormal for a

score of 69 or less, questionable for a score of 70-100, and normal for a score greater than 100. Our data show that low-risk, healthy preterm babies tended to have Premie-Neuro scores on the borderline of the current questionable and normal categories, while sicker high-risk babies tended to score in the middle of the questionable category. Thus, many low-risk, otherwise healthy babies with raw scores just under 100 were classified as neurologically questionable during the course of the study. No infants in this study were classified as abnormal on any assessment using the currently established cut-points.

2.6 Conclusions

There is a need for a clinical neurological assessment tool that may be used in the NICU to identify infants at-risk for developmental delay. The Premie-Neuro is a quick, non-invasive assessment that may be easily administered by an experienced clinician. This study shows that Premie-Neuro scores meet reliability and validity criteria expected in this fragile population and thus may be used as an adjunct to clinical assessment to identify at-risk infants. However, Premie-Neuro classifications do not have adequate reliability and validity and should be interpreted with caution. Modification of Premie-Neuro cut-points may enhance its reliability, validity, and clinical utility as an early screening for developmental disability and delay. Future research is needed to further determine the psychometric properties of the Premie-Neuro and to determine if modification of classification cut-points does, indeed, result in enhanced reliability and validity.

Subtest	Items
Neurologic	<ol style="list-style-type: none"> 1. Arm Recoil 2. Arm Traction 3. Palmar Grasp 4. Plantar Grasp 5. Scarf Sign 6. Popliteal Angle 7. Heel to Ear 8. Movement Type
Movement	<ol style="list-style-type: none"> 1. Tremors 2. Thrashing 3. Facial Grimace 4. Startle 5. Yawn 6. Color Change 7. Arm Movements 8. Leg Movements
Responsiveness	<ol style="list-style-type: none"> 1. Arm Flexion 2. Head Lag 3. Held Sit 4. Posterior Neck 5. Anterior Neck 6. Alert 7. Ventral Suspension 8. Responsiveness

Table 2.1. Premie-Neuro Test Items. The Responsiveness subtest may only be administered to infants \geq 28 weeks PMA and who do not require mechanical ventilation. The Neurologic and Movement subtests may be administered to all preterm infants.

	Sex	GA at birth, weeks	BW, grams	Birth order	Race/ ethnicity
High Risk Group (n=11)	10 male 1 female	25.45 (1.81)	801.64 (264.91)	7 singletons 3 twins 1 triplet	9 White 0 African-American 1 Asian 1 Hispanic
Low Risk Group (n=23)	15 male 8 female	30.61 (3.23)	1602.17 (691.73)	9 singletons 11 twins 3 triplets	18 White 2 African-American 1 Asian 2 Hispanic
Total (n=34)	25 male 9 female	29 (3.67)	1343.18 (696.25)	16 singletons 14 twins 4 triplets	27 White 2 African-American 2 Asian 3 Hispanic

Table 2.2. Description of study participants. Data are mean (standard deviation). GA=Gestation Age; BW=Birth Weight.

	Interrater reliability (n=15)	One-day test-retest reliability (n=30)
Movement and Neurologic subtests only	ICC=0.556 κ =0.400	ICC=0.493 κ =0.429
Movement, Neurologic, and Responsiveness subtests	ICC=0.391* κ =0.143*	ICC=0.592** κ =0.133**

Table 2.3. Intraclass correlation coefficients for interrater and test-retest

reliabilities of Premie-Neuro raw score, and Kappa agreements for interrater and test-retest reliabilities of Premie-Neuro classifications. Interrater and test-retest reliabilities of Premie-Neuro raw scores were fair to moderate for two and three subtests. Interrater and test-retest reliabilities of the Premie-Neuro classifications was fair to moderate for two subtests, but only slight for all three subtests.

***n=14, One infant was on a ventilator at the time of interrater reliability testing and was only administered 2 subtests.**

****n=26, Four infants were on a ventilator at the time of test-retest reliability testing and were only administered 2 subtests.**

Subject	Premie-Neuro test week										Most frequent classification	% of assessments in most frequent classification
	1	2	3	4	5	6	7	8	9	10		
1	Q	N	N	Q	Q	Q	Q				Q	71.43
2	N	Q	N	N	N	Q	N				N	71.43
3	N	N									N	100
4	Q	N	Q								Q	66.67
5	Q	N	N								N	66.67
6	N	N	Q	Q	Q	Q					Q	66.67
7	Q	N	Q	Q	Q	Q					Q	83.33
8	Q	Q									Q	100
9	N	N	N	N	Q	N					N	83.33
10	N	N	N	Q	Q	Q					N/Q	50
11	N	N									N	100
12	Q	Q									Q	100
15	N	Q	N	N	Q	Q	Q				Q	57.14
16	Q	Q	Q	N	Q	Q	Q	Q			Q	87.5
18	Q	N	N	Q	Q	Q	Q	Q	Q	Q	Q	80
20	N	N	Q								N	66.67
24	N	N	N	N	Q	Q					N	66.67
25	Q	Q	N								Q	66.67
26	N	Q									N/Q	50
28	N	N	N								N	100
29	Q	Q	Q	Q							Q	100
30	Q	Q	Q	Q	Q	Q	Q				Q	100
31	Q	Q	Q								Q	100
32	N	Q	Q	Q							Q	75
33	Q	N	Q	Q	Q	Q					Q	83.33
34	N	Q	Q	Q							Q	75
												Mean= 79.52

Table 2.4. Stability of Premie-Neuro classifications across weeks. N=Normal;

Q=Questionable.

	First PN score	Discharge PN Score
High Risk Group (n=11)	2 subtests: 64.36 (7.74)	2 subtests: 58.55 (7.85)
	3 subtests (n=7)*: 95.43 (6.90)	3 subtests: 84.73 (9.93)
Low Risk Group (n=23)	2 subtests (n=15)**: 71.07 (4.27)	2 subtests: 70.52 (4.40)
	3 subtests (n=15)**: 102.93 (5.80)	3 subtests: 100.26 (6.86)
Total (n=34)	2 subtests (n=26)**: 68.23 (6.75)	2 subtests: 66.65 (8.00)
	3 subtests (n=22)***: 100.55 (6.99)	3 subtests: 95.24 (10.75)

Table 2.5. Premie-Neuro scores for high-risk and low-risk infants. Data are mean (standard deviation). PN=Premie-Neuro.

*** Four infants were on ventilators at first PN assessment and were only administered 2 subtests**

****Eight infants were only administered one weekly PN assessment. In that case, the assessment was entered as discharge PN.**

Variable	Neurologic and Movement subtests only	Neurologic Movement, and Responsiveness subtests
PMA	p=0.001*	p=0.000*
Risk Group	p=0.000*	p=0.000*
PMA x Risk Group	p=0.004*	p=0.010*

Table 2.6. Two Factor (Risk Group, PMA) Repeated Measures Analysis of

Variance (n=34). Premie-Neuro total raw score was the dependent variable.

There were significant main effects of both PMA and Risk Group, and a significant interaction between PMA and Risk Group. PMA=post-menstrual age.

***p<0.05**

	Dependent variable: discharge PN total raw score	Dependent variable: discharge PN classification
Predictors in the model	BPD (p=0.013)* CNS lesion (p=0.024)*	BPD (p=.038)* CNS lesion (p=0.999)

Table 2.7. Linear and logistic regression analyses with stepwise elimination.

BPD and CNS lesion were significant predictors of discharge PN raw score, only BPD was a significant predictor of discharge PN classification. Gestational age at birth was not included in either regression model due to lack of significance. BPD=Bronchopulmonary Dysplasia defined as oxygen dependence at 36 weeks PMA; CNS=central nervous system; PN=Premie-Neuro.

***p<0.05**

PMA (weeks)	Number of assessments		
	HR	LR	Total
29	2	1	3
30	2	1	3
31	6	5	11
32	5	6	11
33	5	6	11
34	7	8	15
35	8	16	24
36	9	14	23
37	6	11	17
Total	50	68	118

Table 2.8. Number of full Premie-Neuro assessments (3 subtests) administered at each week PMA. HR=high risk group; LR=low risk group; PMA=Post-menstrual Age.

2.7 Figure Legend

Figure 2.1. Weekly Premie-Neuro scores for HR and LR infants. After 32 weeks PMA, mean scores for LR infants tended to remain stable while scores for the HR infants declined as they approached term adjusted age. LR=Low Risk Group; HR=High Risk Group; PMA=Post-menstrual Age.

Figure 2.2. Weekly Premie-Neuro scores for HR and LR infants. Data collected at 29 and 30 weeks post-menstrual age were removed as there was only one data point in the LR group and two data points in the HR group for each of those weeks. There were ≥ 5 data points for each risk group for post-menstrual weeks 31-37 (Table 2.8).

Figure 2.1

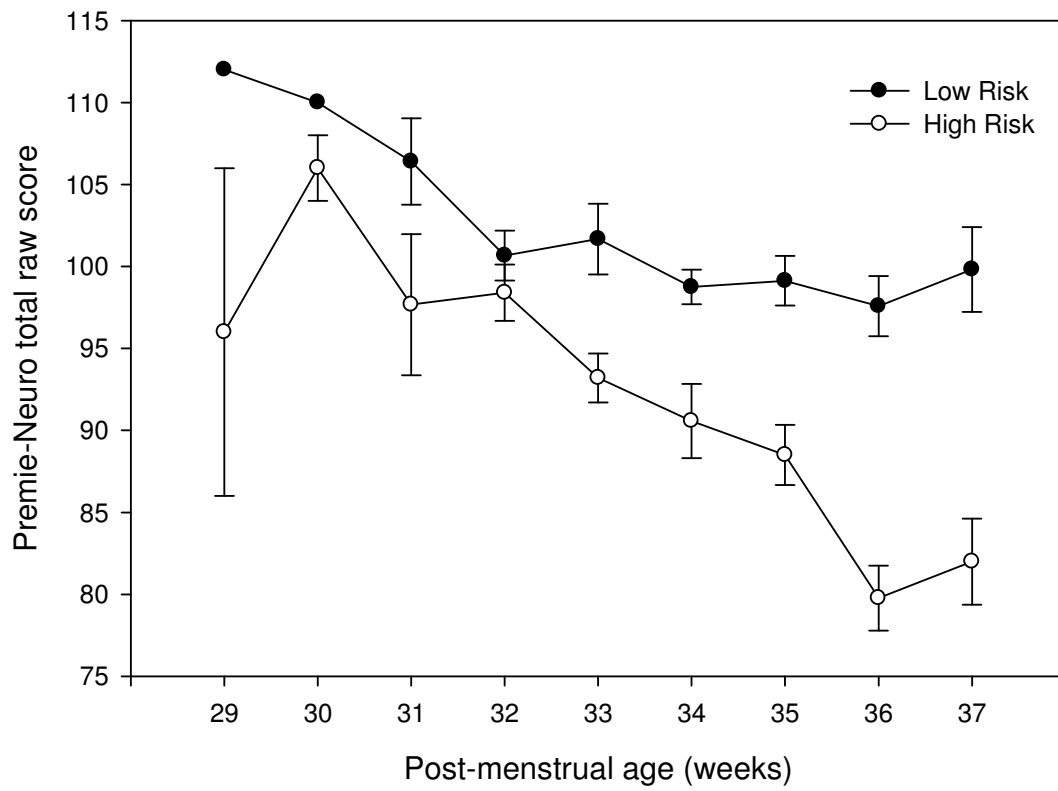
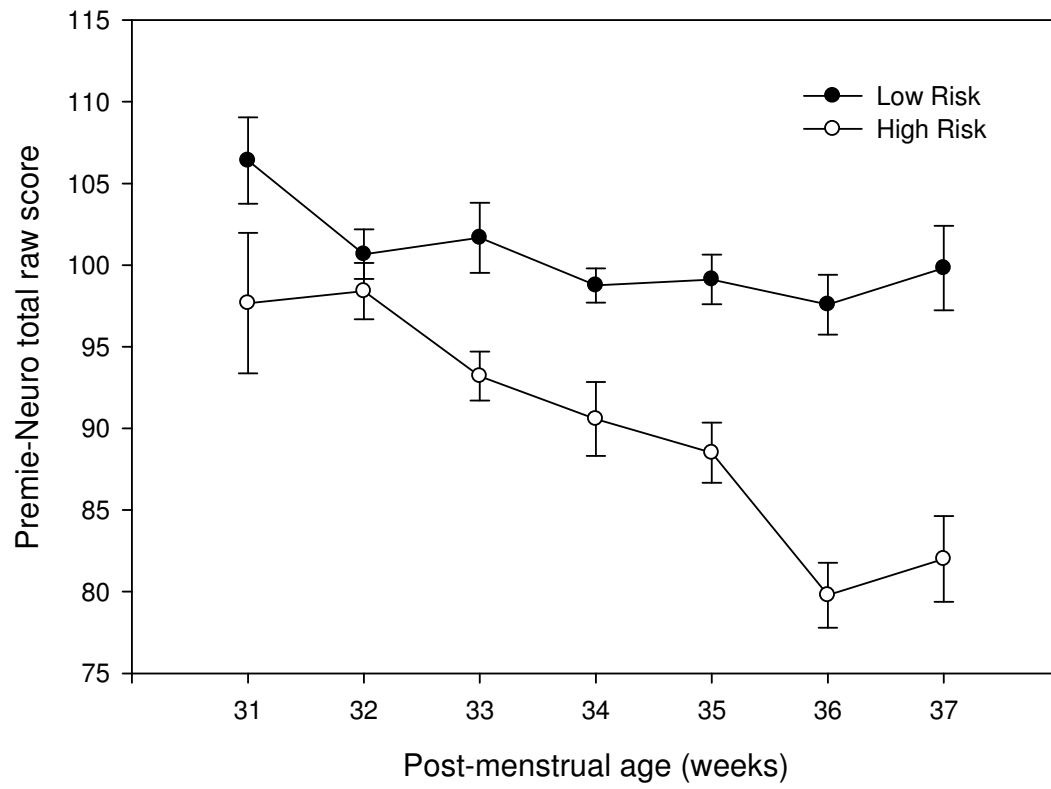


Figure 2.2



Chapter 3

Predictive Validity of the Premie-Neuro at Term and 3 Months Adjusted Ages

3.1 Abstract

Predicting long-term outcomes for the preterm infant is a complex and difficult task, and there are few standardized assessment tools for this population with established predictive validity. In the present study, we set out to determine if performance on the Premie-Neuro—a standardized neurological assessment for preterm infants 23 through 37 weeks post-menstrual age (PMA)—could predict neurodevelopmental outcomes at term and 3 months adjusted age. Thirty-four preterm infants were administered the Premie-Neuro weekly in the NICU through 37 weeks post-menstrual age (PMA) or discharge, whichever occurred first. At 38-42 weeks PMA, infants were assessed using the NeoNeuro. At 3 months adjusted age, infants were assessed using the Infanib and AIMS. Based on their performance, infants were classified as normal or not normal on all assessments based on recommended cut-points. We found that Premie-Neuro raw scores had fair to moderate ability to predict outcomes at term and 3 months adjusted age, and that the predictive value of Premie-Neuro scores improved as infants approached term adjusted age. Premie-Neuro classifications were generally not predictive of NeoNeuro classifications at term or Infanib and AIMS classifications at 3 months adjusted age. These data suggest that Premie-Neuro raw scores may be used by the clinician to identify infants at-risk for neurodevelopmental delays. However, Premie-Neuro classifications should be interpreted cautiously. Future research is needed to assess the role of growth and development on Premie-Neuro score and to determine if modification of raw score cut-points used to classify assessments as normal,

questionable, or abnormal would enhance predictive validity of Premie-Neuro classification.

3.2 Introduction

Prematurity—defined as birth prior to 37 weeks gestational age—is one of the leading causes of infant death and neurodevelopmental disabilities such as cerebral palsy, developmental coordination disorder, learning disabilities, and low IQ (2, 3, 21, 43, 93, 94). Although improvements in medical care since the mid- to late-1990s have led to improved neurodevelopmental outcomes in preterm infants (6, 7), the rate of preterm birth in the United States continues to rise each year (1) and there are disproportionate numbers of preterm children who require specialized care, early intervention services, and special education throughout childhood. For example, in Massachusetts, Clements and colleagues reported that the cost of early intervention services per preterm child was over 2-7 times the amount spent per term child. The largest discrepancy in dollars spent was for physical and occupational therapy, reflecting a focus on motor delay and disability in this population (64). However, research indicates that early intervention does not improve motor outcomes for preterm infants who receive such services simply because they are preterm (9, 10). Rather, early intervention is only proven to be effective for a subset of preterm infants who are identified as at-risk using standardized neurological assessment tools (12, 13). In a time when rates of preterm birth are rising and state and federal budgets are shrinking, it is increasingly important to maximize efficiency and cost-effectiveness of early intervention services for preterm children. Early prediction of neurodevelopmental delay has never been more important.

Prediction of neurodevelopmental outcomes in preterm infants is complicated, and it is particularly difficult to predict individual outcomes. Neuroimaging—often cranial ultrasound or, less frequently, magnetic resonance imaging (MRI)—is often used to identify children with brain lesions as “high risk.” However, children with documented brain lesions often have no functional neurological deficits, while children without evidence of early brain injury sometimes go on to have significant disability (24). Research suggests that the prediction of neurological dysfunction is best when perinatal risk factors, including results of neuroimaging, are considered along with the results of clinical neurological assessments (57, 58). Unfortunately, there are few standardized assessment tools for use in the NICU with established predictive validity, so the clinician must often rely on his or her own clinical expertise to predict outcomes. Clinicians may struggle with the choice of whether to refer groups of preterm infants for early intervention services based on risk factors—resulting in a potential over-utilization of services—versus possibly under-utilizing services while waiting to make those referrals until evidence of delay or disability is undeniable. A predictive standardized neonatal neurological assessment tool would aid the clinician in making prognoses and appropriately referring infants—or not—for developmental follow-up and/or early intervention services.

The Premie-Neuro is a standardized neurological assessment tool for infants 23 through 37 weeks post-menstrual age (PMA). It contains items that measure active and passive tone, reflexes, and behavioral signs of stress. The Premie-Neuro consists of 8 items in each of 3 subtests—Neurological, Movement, and

Responsiveness—for a total of 24 items (78). It requires little training and can be easily administered in 5-10 minutes by the experienced clinician. Internal consistency of the Premie-Neuro is good (14). In Chapter 2, we reported results from our work showing that Premie-Neuro raw scores had fair to moderate interrater and test-retest reliability. Construct validity was also strong as the Premie-Neuro distinguished between groups of infants known to differ in terms of risk for developmental delay and disability. Although preliminary work has shown that the Premie-Neuro may have acceptable predictive validity through 6-8 months adjusted age (79), this has never been formally studied. Thus, the purpose of the present study was to determine the predictive validity of the Premie-Neuro. Specifically, we set out to determine if performance on the Premie-Neuro in the NICU could predict performance on the NeoNeuro, a neurological assessment for young infants (95), at 38-42 weeks PMA. We also examined the relationship between NICU Premie-Neuro performance and performance on the Infanib (90), a neurological test for older infants, and Alberta Infant Motor Scale (96), a test of motor development for older infants, at 3 months adjusted age.

3.3 Methods

Participants

Thirty-four preterm infants were tested with the Premie-Neuro at least once in the NICU and participated in follow-up testing at term and/or 3 months adjusted age. A description of the infants who participated in this study is provided in table 3.1.

Infants were included in the study if they were born prior to 37 weeks gestation, were no more than 37 weeks PMA at the time of the first Premie-Neuro assessment, and if their physician agreed that they were medically stable enough to undergo hands-on clinical assessment. Infants were excluded from the study if they had a known genetic disorder or major congenital anomaly. In accordance with the Human Subjects Committee, each participant's parent or guardian signed an institutionally approved parental consent form.

NICU testing procedure

Premie-Neuro testing began as soon as possible after the infant's physician determined each participant was sufficiently medically stable to undergo hands-on assessment. All Premie-Neuro testing was conducted by a physical therapist in the infant's crib or isolette, and all assessments were completed within the hour prior to each infant's scheduled feeding time when possible. Prior to each test, the infant's physician and/or nurse were contacted to ensure there were no medical complications—such as an acute illness or infection—that would limit the babies' ability to participate in the assessments. Physiological signs of stress—such as heart rate and oxygen saturation—were monitored throughout each testing session, and testing was discontinued and/or canceled and rescheduled if the infant showed significant signs of stress. The Premie-Neuro was administered approximately weekly (once every 5-9 days) through 37 weeks PMA or discharge, whichever occurred first.

In accordance with administration guidelines, infants who were 28 weeks PMA or younger or who were on a ventilator were only administered the Neurologic and Movement subtests of the Premie-Neuro, a total of 16 items. Infants older than 28 weeks PMA and not on a ventilator were administered all 24 items in the 3 subtests. Each item was assigned a score of one, three, or five based on the infant's gestational age and performance on the item. Item scores were tallied and each infant's assessment was classified as neurologically normal, questionable, or abnormal based on raw score cut-points provided in *The Premie-Neuro Examination Instruction Manual* (78).

Follow-up testing at term adjusted age

At 38-42 weeks PMA, the NeoNeuro was administered by a physical therapist. If the infant remained in the NICU at term adjusted age, testing was conducted in the NICU at bedside. If the infant was discharged from the NICU, testing was conducted in the home or the NICU follow-up clinic depending on parent preference.

The NeoNeuro consists of 32 items assessing tone and movement, reflexes, and neurobehavior in infants from 38 weeks PMA to 16 weeks adjusted age. Each item is assigned a score of 1, 3, or 5 based upon the infant's adjusted age and performance on the item. Individual item scores are tallied for a total score, and the assessment is classified as normal, mildly abnormal, moderately abnormal, or severely abnormal based on raw score cut-points recommended by the NeoNeuro

authors (95). Reliability and validity of the NeoNeuro has been tested. Internal consistency is 0.80 and test-retest reliability over one week is 0.73. Construct validity of the NeoNeuro has been demonstrated by its ability to distinguish between infants who are neurologically normal, suspect, or abnormal (65, 95). We chose to assess infants at term adjusted age using the NeoNeuro because the items it includes—assessing active and passive tone, reflexes, and neurobehavior—measure a similar construct as the Premie-Neuro.

Follow-up testing at 3 months adjusted age

At approximately 3 months adjusted age, infants were assessed by a physical therapist using two tools: the Alberta Infant Motor Scale (AIMS) and the Infanib. No infants remained in the NICU at 3 months adjusted age, so all testing was conducted in the infant's home or the NICU follow-up clinic according to caregiver preference.

The Infanib is a criterion-referenced neuromotor assessment that may be administered to infants age birth to 18 months. The Infanib consists of 20 items, 15 of which are appropriate to administer to a child who is less than 4 months adjusted age. All 15 of those items were administered to each participant, and each infant was given a score of 1, 3, or 5 based on performance on the item. Item scores were tallied for a total score, and each assessment was categorized as normal, transient, or abnormal based on raw score cut-points provided in the administration manual (90). Previous studies have established the reliability and validity of the Infanib for use in

research and clinical practice, determining that the interrater reliability was 0.97 and test-retest reliability was 0.95 (90). We chose to administer the Infanib at 3 months adjusted age because—like the Premie-Neuro and NeoNeuro—the Infanib assesses tone and posture. Thus, we believe the Infanib measures a similar construct as the Premie-Neuro and NeoNeuro.

The AIMS is a norm-referenced assessment that can be administered on infants 0-18 months (96). The AIMS contains four subscales: prone, supine, sit, and stand. The AIMS was administered to each participant in its entirety by observing motor performance in each of these positions. Total raw score was tallied and a percentile rank was calculated based on the infant's adjusted age. Each infant's assessment was classified as normal if gross motor skills fell at or above the 10th percentile rank and abnormal if motor skills fell below the 10th percentile rank (97). The psychometric properties of the AIMS have been well-documented. Using the 10th percentile as a cut-point, the AIMS had a specificity of 81.7% and a sensitivity of 77.3% when administered on 4-month-old infants. Interrater reliability of the AIMS was >0.95 and test-retest reliability ranged from 0.86-0.99. Concurrent validity of the AIMS with other established motor evaluations ranged from 0.84 to 0.99 (68, 96). The AIMS differs from the Premie-Neuro, NeoNeuro, and Infanib in that it is a measure of gross motor development and does not focus on specific neurological signs such as primitive reflexes and muscle tone. In the present study, we chose to administer the AIMS at 3 months adjusted age to determine if the predictive abilities

of the Premie-Neuro extended beyond neurological signs to functional movement skills

Data analysis

In order to assess the predictive validity of the Premie-Neuro raw scores, Pearson correlations were calculated to determine the relationship between Premie-Neuro scores in the NICU and NeoNeuro score at term and Infanib and AIMS scores at 3 months adjusted age. In the present study, we chose to test correlations associated with raw scores from both the first Premie-Neuro test and the final Premie-Neuro test administered before discharge from the NICU. During the course of our study, no Premie-Neuro assessment was classified as abnormal. Thus, to examine predictive validity of Premie-Neuro classifications, we did not change the author recommendations for classifying an assessment as “normal,” but we did collapse the author-recommended questionable and abnormal classifications into one “not normal” classification. Similarly, we collapsed author-recommended NeoNeuro and Infanib classifications into “normal” and “not normal.” Recommended AIMS classifications were binary, so no modification of classifications was necessary for data analysis. We compared Premie-Neuro classifications in the NICU to NeoNeuro classifications at term and Infanib and AIMS classifications at 3 months adjusted age using Fisher’s exact test. Again, we chose to look at only classifications obtained at the first and discharge Premie-Neuro assessments.

Eight infants were only administered one Premie-Neuro assessment while in the NICU due to discharge before a second test could be administered. These scores and classifications were included in analyses for the discharge Premie-Neuro assessment. Four infants were on a ventilator at the time of the first Premie-Neuro assessment. Per testing guidelines, those infants were administered only the Neurologic and Movement subtests at their first assessment. Thus, of the 34 infants included in our sample, scores for all 3 subtests were recorded and analyzed for 22 infants and scores for 2 subtests were obtained for 26 infants at the first Premie-Neuro assessment. All infants were off the ventilator at the time of the discharge Premie-Neuro assessment, so scores for 2 and 3 subtests were recorded and analyzed for all 34 infants for the discharge Premie-Neuro assessment.

Of the original 34 subjects, one infant's family could not be contacted for term follow-up. Thus, 33 infants were included in analyses for predictive validity at term adjusted age. At 3 months adjusted age, five infants' families could not be reached, three infants' families had relocated, and one infant was deceased. Thus, 25 infants were included in analyses for predictive validity at 3 months adjusted age. Analyses were performed with SPSS 16.0*. Significance was set at $p < 0.05$. All data are reported as mean \pm standard deviation.

lxvii

* SPSS, 233 S. Wacker Drive, Chicago, Illinois, 60606

3.4 Results

At the first Premie-Neuro assessment, infants ranged in age from 29 to 36 weeks PMA, with a mean PMA at first Premie-Neuro assessment of 32.35 weeks (± 2.48 weeks). At the discharge Premie-Neuro assessment, infants ranged in age from 32 to 37 weeks PMA (mean PMA 36.0 ± 1.18 weeks). Mean scores for 2 and 3 subtests at the first and discharge Premie-Neuro assessments are summarized in Table 3.2.

Predictive validity of the Premie-Neuro at term adjusted age

Pearson correlations between NeoNeuro scores at term and Premie-Neuro raw scores at the first and discharge Premie-Neuro assessments for infants who were administered both two and three subtests ranged from 0.489-0.700. All correlations were statistically significant (Table 3.3, Figure 3.1.a).

There was no significant association between the first Premie-Neuro classifications and NeoNeuro classifications at term ($p=0.411$). However, there was a significant association between the discharge Premie-Neuro classifications and NeoNeuro classifications at term ($p=0.003$, Table 3.4). The first Premie-Neuro classifications were 59% sensitive and 63% specific for predicting NeoNeuro classifications at term. Sensitivity and specificity of the discharge Premie-Neuro classifications for predicting NeoNeuro classifications at term was 85% and 69%, respectively (Table 3.5).

Predictive validity of the Premie-Neuro at 3 months adjusted age

Table 3.3 and Figures 3.1.b and 3.1.c describe Pearson correlations between Premie-Neuro raw scores and Infanib and AIMS raw scores at 3 months adjusted age. All Pearson correlations between the discharge Premie-Neuro scores for both 2 and 3 subtests and Infanib and AIMS scores at three months adjusted age were statistically significant, ranging from 0.446-0.573. The first Premie-Neuro scores for infants administered 2 or 3 subtests were not significantly correlated with Infanib or AIMS scores at 3 months adjusted age.

While there was a significant correlation between the raw scores of the discharge Premie-Neuro assessments and the Infanib and AIMS raw scores at 3 months adjusted age, the ability of the Premie-Neuro to predict neurologically “normal” versus “not normal” classifications of babies at 3 months adjusted age was not good. There was not a statistically significant relationship between the first or the discharge Premie-Neuro classifications and Infanib or AIMS classifications at 3 months adjusted age (Table 3.3). The first Premie-Neuro classifications were 75% sensitive and 53% specific in predicting AIMS classifications and 50% sensitive and 40% specific in predicting Infanib classifications at 3 months adjusted age. Sensitivity of the discharge Premie-Neuro classifications was 100% and specificity was 57% for predicting AIMS classifications. Sensitivity and specificity of the discharge Premie-Neuro classifications in predicting Infanib classifications were 61% and 71%, respectively (Table 3.5).

3.5 Discussion

Despite great improvements in rates of major medical complications in preterm infants, there is no question that these babies are at greater risk for developmental disability and delay than their term counterparts. Research indicates that early intervention services may improve neuromotor outcomes, but only in a subset of preterm babies who are identified as at-risk (9, 10, 12, 13). Thus, early detection of at-risk infants is critically important so that clinicians may make appropriate recommendations for referring or not referring preterm babies for specialized early intervention services such as physical or occupational therapy. There is growing consensus that standardized neurological testing should be considered in combination with clinical assessment, medical risk factors, and medical testing to improve the prediction of poor outcomes in preterm babies (57, 58). In the present study, we set out to determine whether one such standardized test—the Premie-Neuro—had acceptable predictive validity to be used to identify at-risk infants in the NICU. Our findings suggested that Premie-Neuro raw scores were predictive of performance on the NeoNeuro at term and the Infanib and AIMS at 3 months adjusted age. However, Premie-Neuro classifications did not seem to predict the classifications of other standardized tests at term and 3 months adjusted age.

As expected, the discharge Premie-Neuro raw score was a stronger predictor of outcomes at both term and 3 months adjusted age than the first Premie-Neuro score. This may be due in part to the fact that neurological development is often

confounded with sickness at early preterm ages. Although sicker babies with more medical complications do tend to have worse neurological outcomes, this is not necessarily a causal relationship and poor performance by a very sick preterm infant may not be reflective of his or her true developmental capabilities. By the discharge Premie-Neuro assessment, all infants in the present study were off of a ventilator and were sufficiently medically stable to be near discharge from the NICU. As such, the discharge Premie-Neuro assessment was probably much more reflective of true infant neurological function than the first Premie-Neuro assessment. Furthermore, the discharge Premie-Neuro assessment was temporally closer to term and 3-month assessments, which likely resulted in increased correlations between the discharge Premie-Neuro scores and scores at follow-up testing. During the course of this study, Premie-Neuro raw scores seemed to have a floor effect. In other words, at very young preterm ages, some Premie-Neuro items—particularly those included in the Responsiveness subtest—did not seem to be sensitive enough to detect neurodevelopmental delays and differentiate between normal and abnormal responses. For example, an item in the Responsiveness subtest examines whether the baby can hold his head up in supported sitting. The inability to hold the head upright in this position is considered a normal response through 33 week PMA. Thus, for many babies who would later struggle with head control and score abnormally on this item, the early response of the head remaining forward or backward was considered normal and difficulty with this item was often not detected until the baby could not hold his head upright at 34-37 weeks PMA. Because our data show quite clearly that

the predictive value of Premie-Neuro score improves with age, further research is needed to determine what other factors (such as respiratory status, illness, growth) may affect the reliability and validity of the Premie-Neuro scores, particularly at early preterm ages. Furthermore, it would be beneficial to determine if there is a certain age or stage of development when the Premie-Neuro scores become significantly more reliable or valid, as this may represent a window of opportunity when Premie-Neuro testing and early detection may be most valuable.

According to Portney and Watkins (98), in the behavioral sciences, lower correlations—those that may not necessarily be statistically significant—may be used as “evidence of functional useful relationships.” Although they caution that rigid cut-points should not be used, they state that correlations between 0.25-0.50 may show fair relationships and correlations 0.51-0.75 may show moderate relationships. Using these criteria, correlations in this study revealed fair to moderate, statistically significant relationships between both the first and discharge Premie-Neuro raw scores in the NICU and NeoNeuro raw scores at term. The discharge Premie-Neuro raw scores also had fair to moderate, statistically significant relationships with Infanib and AIMS scores at 3 months adjusted age. According to the Portney and Watkins criteria, the first Premie-Neuro scores had a fair relationship with Infanib scores at 3 months adjusted age, although this was not a statistically significant relationship. The first Premie-Neuro scores had little to no relationship with AIMS scores at 3 months adjusted age. This discrepancy may exist because the Premie-Neuro, NeoNeuro, and Infanib assess a similar construct (neurological function) while the AIMS assesses a

somewhat different construct (motor development). Another possible explanation is that neurological development at early preterm ages—as measured by the Premie-Neuro—simply may not be as reflective of future motor development as neurological function at preterm ages closer to term.

Although the predictive validity of the Premie-Neuro raw score was generally quite good, Premie-Neuro classifications had poor predictive value overall. The discharge Premie-Neuro classifications were only predictive of NeoNeuro classifications at term and was not predictive of classifications at 3 months adjusted age. The first Premie-Neuro classifications were not predictive of classifications at term or 3 months adjusted age. One limitation of the present study was the relatively small, homogenous sample of infants. As stated previously, during the course of our study, no infant tested in the “abnormal” classification on any Premie-Neuro assessment. Because of this, we chose to collapse classifications of all assessments into two categories—normal and not normal—for data analyses. Perhaps a different method of collapsing categories, such as collapsing normal and questionable Premie-Neuro classifications into one group and abnormal classifications into a second group—would have enhanced predictive validity of Premie-Neuro classifications. Unfortunately, we were unable to do this for our sample because of the lack of infants who tested as abnormal in the study. A larger, more heterogeneous sample of infants may have allowed analyses that would have been able to capture the subtle variations between infants in different classifications. However, even in a large sample, we would still expect very few babies to test in the

abnormal range and many of those babies scoring in the abnormal range of the Premie-Neuro may be too fragile to test at early preterm ages. Thus, the sample tested here appears to represent a typical group of babies that would be available to the clinician for assessment.

Another possible explanation for the poor predictive validity of Premie-Neuro classifications—despite good predictive validity of the Premie-Neuro raw scores—is that the raw score cut-points set by the Premie-Neuro authors during the construction of the test may need to be adjusted. Future studies are needed to determine if adjusting the raw score cut-points and/or modifying the classification system would enhance the predictive validity of the Premie-Neuro classifications and improve the use of the Premie-Neuro as a clinical assessment tool.

3.6 Conclusions

Prediction of outcomes in preterm infants is difficult, even for an experienced clinician, and there are currently few standardized tools available for clinicians to identify at-risk infants and differentiate between infants with neurodevelopmental delays that are mild or severe, transient or persistent. In the current study, we found that the Premie-Neuro raw scores had fair to moderate predictive validity. Although the Premie-Neuro raw scores obtained early in the infants' NICU stays had some predictive value, the discharge Premie-Neuro scores were the strongest predictors of outcomes at term and 3 months adjusted age. The predictive value of the Premie-Neuro classifications at term and 3 months adjusted age was generally poor.

Clinicians should interpret Premie-Neuro classifications with caution. Our data suggest that Premie-Neuro raw scores, particularly at older preterm ages, are predictive of outcomes at term and 3 months adjusted age, and may be used to identify at-risk infants in the NICU. Modification of Premie-Neuro raw score cut-points for classifying assessments as normal, questionable, or abnormal may enhance the predictive value of Premie-Neuro classifications.

	Sex	GA at birth, wks	BW, grams	Birth order	Race/ ethnicity
Total (n=34)	25 male 9 female	29 (3.67)	1343.18 (696.25)	16 singletons 14 twins 4 triplets	27 White 2 African- American 2 Asian 3 Hispanic

Table 3.1. Description of study participants. Data are mean (standard deviation). GA=Gestational Age; BW=Birth Weight.

	First PN score (n=26*)	Discharge PN Score (n=34)
2 subtests	68.23 (6.75)	66.65 (8.00)
3 subtests	100.55 (6.99)**	95.24 (10.75)

Table 3.2. Summary of raw scores at the first and discharge Premie-Neuro assessments. Data are mean (standard deviation). Raw score cut-points for classifications based on 3 subtests: ≥ 100 normal, 70-99 questionable, < 70 abnormal. Raw score cut-points for classifications based on 2 subtests: ≥ 70 normal, 50-69 questionable, < 50 abnormal (78). PN=Premie-Neuro.

***Eight infants were only administered one weekly PN assessment. In that case, the assessment was entered as discharge PN.**

*** *Four infants were on ventilators at first PN assessment and were only administered 2 subtests (n=22).**

	NeoNeuro (n=33)	Infanib (n=25)	AIMS (n=25)
First Premie-Neuro **			
2 subtests	0.489 (p=0.013)*	0.455 (p=0.05)	0.178 (p=0.465)
3 subtests [#]	0.671 (p=0.001)*	0.390 (p=0.121)	-0.042 (p=0.874)
Discharge Premie-Neuro			
2 subtests	0.667 (p=0.000)*	0.446 (p=0.025)*	0.460 (p=0.021)*
3subtests	0.700 (p=0.000)*	0.537 (p=0.006)*	0.525 (p=0.007)*

Table 3.3. Pearson correlation coefficients for first and discharge Premie-Neuro

scores for two and three subtests and NeoNeuro scores at term and Infanib and AIMS scores at three months adjusted age. The discharge Premie-Neuro raw scores were significantly correlated with raw scores obtained using the NeoNeuro at term and the Infanib and AIMS at 3 months adjusted age. The first Premie-Neuro raw scores were significantly correlated with NeoNeuro raw scores at term, but not with Infanib and AIMS raw scores at 3 months adjusted age. AIMS=Alberta Infant Motor Scale.

*p<0.05

** Eight infants were only administered one weekly PN assessment. In that case, the assessment was entered as discharge PN (n=26).

[#]Four infants were on ventilators at first PN assessment and were only administered 2 subtests (n=22).

	First PN	Discharge PN
NeoNeuro	p=0.411	p=0.003*
Infanib	p=1.000	p=0.202
AIMS	p=0.582	p=0.096

Table 3.4. Fisher's exact test for the first and discharge Premie-Neuro classifications and NeoNeuro classifications at term and Infanib and AIMS classifications at 3 months adjusted age. The only statistically significant relationship was between the discharge Premie-Neuro classifications and NeoNeuro classifications at term adjusted age. AIMS=Alberta Infant Motor Scale; PN=Premie-Neuro.

***p<0.05**

	Discharge PN Sensitivity	Discharge PN Specificity
NeoNeuro (term)	85%	69%
AIMS (3 months AA)	100%	57%
Infanib (3 months AA)	61%	71%

Table 3.5. Sensitivity and specificity of the discharge Premie-Neuro

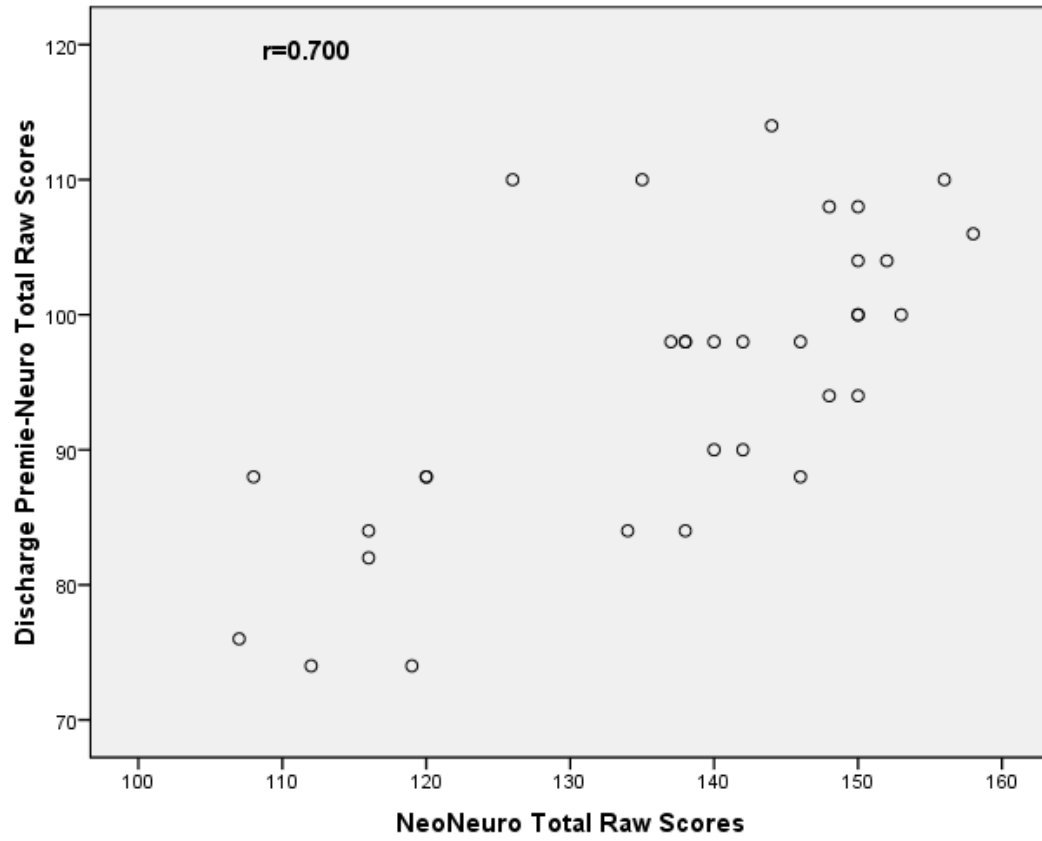
classifications for predicting NeoNeuro classifications at term or Infanib and AIMS classifications at 3 months adjusted age. PN=Premie-Neuro, AIMS=Alberta Infant Motor Scale

3.7 Figure Legend

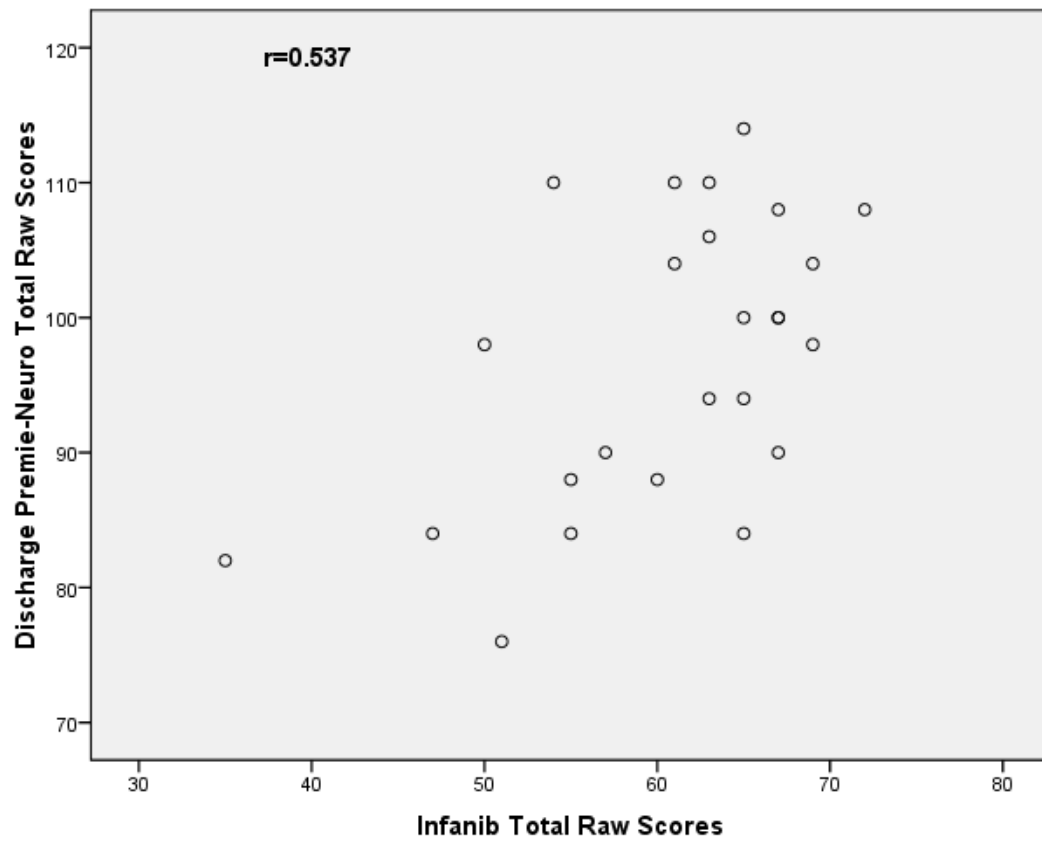
Figure 3.1. Scatterplot showing Pearson correlations between the discharge Premie-Neuro raw scores and (a) NeoNeuro raw scores at term, (b) Infanib raw scores, and (c) AIMS raw scores at 3 months adjusted age. All correlations were significant at $p < 0.05$. AIMS=Alberta Infant Motor Scale.

Figure 3.1

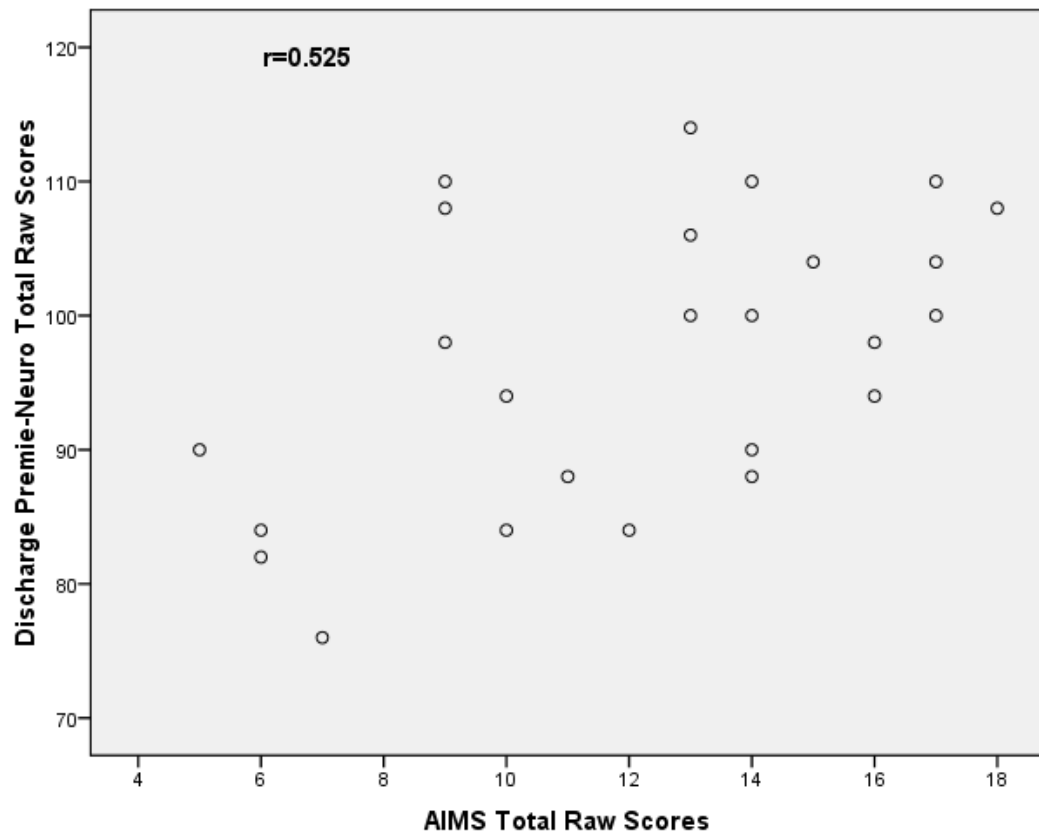
a.



b.



c.



Chapter 4

**Modifying raw score cut-points improves reliability, validity, and predictive
value of Premie-Neuro classifications**

4.1 Abstract

The Premie-Neuro is a new test of neurological development for preterm infants. In Chapters 2 and 3, we showed that Premie-Neuro raw scores had acceptable reliability, could differentiate between groups of infants known to differ in terms of risk for delay and disability, and were predictive of outcomes at term and 3 months adjusted age. However, Premie-Neuro classifications using raw score cut-points recommended in *The Premie-Neuro Examination Instruction Manual* (78) to classify assessments as normal (raw scores ≥ 100) and not normal (questionable and abnormal, raw scores <100) were not reliable and valid. In the present study, we examined whether modifying the raw score cut-points for classifying Premie-Neuro assessments as normal, questionable, or abnormal would improve reliability, validity, sensitivity, and specificity of the tool. Thirty-four preterm infants were administered the Premie-Neuro at least once per week in the neonatal intensive care unit (NICU) through 37 weeks post-menstrual age (PMA) or discharge, whichever occurred first. At 38-42 weeks PMA, infants were assessed using the NeoNeuro, a reliable and valid standardized test for term newborns up to 16 weeks adjusted age. At 3 months adjusted age, infants were assessed using the Infanib and Alberta Infant Motor Scale (AIMS), two reliable and valid standardized assessments for infants up to 18 months adjusted age. Data were analyzed to determine (1) if mean scores obtained on discharge Premie-Neuro assessments were different for infants who later scored in the abnormal versus normal ranges on term and 3-month follow-up assessments and (2) if changing the raw score cut-point used classify an assessment as abnormal enhanced

reliability and validity of Premie-Neuro classifications. We found that, after 34 weeks PMA, mean Premie-Neuro scores for infants with abnormal follow-up were less than 90. Infants with discharge Premie-Neuro raw scores less than 90 performed significantly worse on term and 3-month follow-up assessments than infants with scores 100 or greater. Using a raw score cut-point of <90, rather than the currently recommended cut-point of <100, for classifying an assessment as not normal enhanced test-retest reliability, stability, construct validity, and predictive validity of the Premie-Neuro classifications. As expected, using the raw score cut-point of <90 improved specificity of detecting neurological abnormalities at term and 3 months adjusted age, but resulted in much lower sensitivity. Based on these data, we suggest that the clinician may use a Premie-Neuro raw score cut-point of <90 to more reliably identify at-risk infants who may benefit from early intervention services and/or close follow-up. However, we caution that using a raw score cut-point of <90 may fail to identify a handful of infants who will later demonstrate neurodevelopmental abnormalities at term or 3 months adjusted age. Thus, we suggest that infants with Premie-Neuro raw scores of 90-99 be classified as questionable, meaning that these babies should be re-tested or monitored closely to ensure that neurological abnormalities do not emerge.

4.2 Introduction

The rate of preterm birth has increased steadily in the United States over the last several decades. Currently, over 12% of all births occur prior to 37 weeks completed gestation (1). Recent advances in neonatal care have resulted in improved developmental outcomes for preterm babies (6, 7). However, there continue to be large numbers of children with disabilities and delays resulting from prematurity, and there are few standardized assessment tools available for the neonatal clinician to identify preterm infants at-risk for such disabilities (2-5, 36). The Premie-Neuro is a new standardized clinical neurological assessment tool for preterm infants (14). It consists of 3 subscales, each containing 8 items. The Neurologic subscale assesses primitive reflexes and passive tone, the Movement subscale examines behavioral signs of stress, and the Responsiveness subscale assesses active tone and responsiveness to handling. Because items included in the Responsiveness subscale require changing the infant's position—and may therefore be stressful for the very young or very sick preterm infant—this subscale may only be administered to infants who are older than 28 weeks PMA and who are not ventilator-dependent. Based on the infant's age and performance, he is given a score on each item. Item scores are tallied for a total raw score, and the raw score is converted to a classification of normal, questionable, or abnormal based on raw score cut-points recommended in *The Premie-Neuro Examination Instruction Manual* (78). The Premie-Neuro is unique in that it can be administered in less than 10 minutes without removing the infant from his crib or isolette and without interfering with electronic monitoring or

mechanical ventilation, and may be given to preterm infants as young as 23 weeks PMA (14, 78). These characteristics make the Premie-Neuro a promising tool for serial clinical neurological assessment in the NICU, providing clinicians with a quick, objective criterion for identifying at-risk preterm infants who may require extra supports and services and/or close developmental follow-up.

In Chapters 2 and 3, we reported that Premie-Neuro raw scores had acceptable reliability, could differentiate between infants at high- and low-risk for poor neurological outcomes, and were predictive of outcomes at term and 3 months adjusted age. However, Premie-Neuro classifications (normal, questionable, and abnormal) based on raw score cut-points recommended by the authors of the Premie-Neuro were not reliable and valid. In the present study, we retrospectively analyzed predictive validity data originally presented in Chapter 3 to determine whether there was a significant difference in discharge Premie-Neuro raw scores in infants who performed in the abnormal versus normal range at term and/or 3 month follow-up testing. Using that data, we identified a possible new raw score cut-point for classifying an abnormal assessment and re-analyzed data originally presented in Chapters 2 and 3 to determine if using the new cut-point resulted in improved reliability and validity of the Premie-Neuro.

4.3 Methods

Participants

Thirty-four preterm infants (mean gestational age at birth 29 ± 3.7 SD weeks, birth weight 1343 ± 696 SD grams) participated in this study. Infants were included in the study if they were born prior to 37 weeks gestation and were no more than 37 weeks PMA at the time of the first Premie-Neuro assessment. Infants with a known genetic disorder or major congenital anomaly were excluded from the study. Prior to participation in the study, each participant's parent or guardian signed an institutionally approved parental consent form.

Premie-Neuro testing

Once the infant's physician determined he or she was medically stable enough to participate, Premie-Neuro testing began. All Premie-Neuro tests were administered by a physical therapist. Testing was conducted in the infant's crib or isolette and was scheduled within one hour prior to the infant's regular feeding time when possible. Heart rate and oxygen saturation were monitored during each assessment, and testing was discontinued and/or canceled and rescheduled if the infant showed significant signs of stress during the testing, was acutely ill the day of testing, or if the infant or examiner were unavailable during the scheduled testing time.

During the first week of testing, the Premie-Neuro was administered twice, no more than 72 hours apart. Premie-Neuro testing continued approximately weekly

(once every 5-9 days) by the same examiner through 37 weeks PMA or discharge, whichever occurred first.

In accordance with the *Premie-Neuro Examination Instruction Manual*, each Premie-Neuro item was assigned a score of 1, 3, or 5 based on the infant's gestational age and performance on the item. Item scores were tallied for a total raw score (14).

Follow-up testing at term adjusted age

At 38-42 weeks PMA, the NeoNeuro was administered. The NeoNeuro is a reliable and valid standardized test that assesses active and passive tone, reflexes, and neurobehavior in infants 38 weeks PMA through 16 weeks adjusted age (65, 95). If the infant remained in the NICU at term adjusted age, testing was conducted at bedside. If the infant was discharged from the NICU, testing was conducted in the home or the NICU follow-up clinic, depending on parent or caregiver preference. In accordance with administration guidelines, each NeoNeuro item was assigned a score of 1, 3, or 5 based upon the infant's adjusted age and performance on the item. Item scores were tallied for a total raw score, and infants were classified as normal, mildly abnormal, moderately abnormal, or severely abnormal based on cut-points recommended by the NeoNeuro authors (65, 95). For this study, NeoNeuro assessments classified as normal or mildly abnormal were considered normal at term follow-up, and assessments classified as moderately or severely abnormal were considered abnormal at term follow-up.

Follow-up testing at 3 months adjusted age

Two reliable and valid standardized assessment tools were used to examine study participants at 3 months adjusted age: the Infanib (90) and the Alberta Infant Motor Scale (AIMS) (90, 96). No infants remained in the NICU at 3 months adjusted age, so testing was conducted in the infants' homes or the NICU follow-up clinic, according to parent or caregiver preference.

The Infanib is a criterion-referenced standardized test that assesses tone, reflexes, and posture. Infants in the present study were tested with the Infanib according to administration guidelines and were classified as neurologically normal, transient, or abnormal based on total raw score cut-points recommended by the Infanib authors (90). The AIMS is a reliable and valid, norm-referenced test of motor development for infants 0-18 months adjusted age (68, 96). In this study, the AIMS was administered in a standardized fashion as described in the administration manual (96). Based on total raw score, each infant was assigned a percentile rank. Using the criteria recommended by Darrah and colleagues, infants with a percentile rank at or above the 10th percentile were classified as normal, while infants who scored <10th percentile were classified as abnormal (97). For this study, infants with Infanib assessments classified as abnormal OR an AIMS scores <10th percentile were considered abnormal at 3-month follow-up. Infants with Infanib assessments classified as transient or normal AND an AIMS score \geq 10th percentile were considered normal at 3-month follow-up.

Demographic and medical factors

At the time of enrollment in the study, the infant's birth weight, sex, and gestational age at birth were obtained by reviewing the infant's medical chart. At discharge, the infant's medical chart was reviewed to obtain the following information: blood pH history, seizure activity, diagnosed central nervous system injury including IVH/PVH and periventricular leukomalacia (PVL), respiratory history (including need for mechanical ventilation), and history of hypoglycemia. These data were used to calculate a Neurobiologic Risk Score (NBRS) for each infant (86). Based on the criteria described by Brazzini and colleagues, infants with a discharge NBRS of <5 were considered "low risk" (Low Risk Group, LR) and infants with a discharge NBRS of ≥ 5 were considered "intermediate/high risk" (High Risk Group, HR) for neurological injury and poor neurodevelopmental outcomes (85).

Data analysis

Only scores obtained for the full Premie-Neuro—all three subtests—were included in data analyses for this study. We performed retrospective analyses of Premie-Neuro assessment results obtained in the NICU to compare Premie-Neuro raw scores for infants with normal follow-up testing at term and 3 months adjusted age versus infants with abnormal follow-up testing at term or 3 months adjusted age. Based on those data, we chose to analyze test-retest reliability, stability, construct validity, and predictive validity of Premie-Neuro classifications using a new raw

score cut-point of <90 to classify a Premie-Neuro assessment as questionable/abnormal. We compared this to reliability and validity analyses presented in Chapter 2 and 3 (using the author-recommended raw score cut-point of 100) to determine whether modification of raw-score cut-points improved test-retest reliability, stability, validity, sensitivity, and specificity and Premie-Neuro classifications.

Kappa agreements were used to analyze one-day test-retest reliability of Premie-Neuro classifications. To examine the stability of Premie-Neuro classifications, we determined the most frequent classification for each infant, and calculated the percentage of assessments that each infant tested in his or her most frequent classification.

In Chapters 2 and 3, we found that the last Premie-Neuro scores obtained prior to discharge from the NICU had significantly better construct and predictive validity than the first Premie-Neuro scores. Thus, for this study, we used only discharge Premie-Neuro classifications to analyze construct and predictive validity, sensitivity and specificity. If an infant was only tested once in the NICU, that score was considered the discharge Premie-Neuro assessment. Fisher's exact test was used to determine if a larger proportion of low-risk infants tested in the normal classification compared to high-risk infants. To examine predictive validity of Premie-Neuro classifications, we collapsed classifications of all follow-up assessments into two classifications—normal and not normal. NeoNeuro assessments in the normal range were classified as normal, while assessments in the mildly, moderately, and severely

abnormal ranges were classified as not normal. Similarly Infanib assessments in the normal range were classified as normal while assessments in the transient or abnormal range were classified as not normal. Finally, AIMS scores <10th percentile were classified as not normal while scores above the 10th percentile were classified as normal. We compared discharge Premie-Neuro classifications to NeoNeuro classifications at term and Infanib and AIMS classifications at 3 months adjusted age using Fisher's exact test. In order to assess the relationship between risk factors (GA at birth, CNS injury, and oxygen dependence at 36 weeks) and discharge Premie-Neuro classifications, logistic regression was used. Analyses were performed with SPSS 16.0*. Significance was set at $p < 0.05$.

4.4 Results

Premie-Neuro scores in infants who were abnormal versus normal at follow-up

Table 4.1 provides descriptive information about the infants who participated in this study. One infant could not be tested for term or 3-month follow-up, and was therefore not included in either the normal or abnormal follow-up groups. At early gestational ages, there was little difference in Premie-Neuro scores between infants who were normal versus abnormal at follow-up. However, beginning at approximately 33-34 weeks PMA, scores for infants who were normal at follow-up remained consistent, while scores for infants who were abnormal at follow-up worsened (Figure 4.1). Mean Premie-Neuro scores from 35-37 weeks PMA for the

xcv—————

* SPSS, 233 S. Wacker Drive, Chicago, Illinois, 60606

normal follow-up group ranged from 96.3-99.3, while raw scores for the abnormal follow-up group ranged from 80-87.7 (Table 4.2). Premie-Neuro raw scores for individual subjects are plotted in Figure 4.2.a. and 4.2.b.

Figure 4.3 shows performance at term (4.3.a) and 3-month (4.3.b and 4.3.c) follow-up for infants with discharge Premie-Neuro raw scores <90, 90-99, and >99. Analysis of variance revealed that infants with discharge Premie-Neuro raw scores <90 performed significantly worse on NeoNeuro assessments at term ($p=0.000$) and Infanib and AIMS assessments ($p=0.001$ and $p=0.009$, respectively) at 3 months adjusted age compared to infants with discharge Premie-Neuro scores >99. Compared to infants with discharge Premie-Neuro scores ranging from 90-99, infants with discharge Premie-Neuro scores <90 performed significantly worse on NeoNeuro assessments at term ($p=0.000$) and Infanib assessments at 3 months adjusted age ($p=0.025$) but not on AIMS assessments at 3 months adjusted age. These data indicated that there were significant differences in developmental and neurological outcomes at term and 3 months adjusted age in infants with discharge Premie-Neuro scores <90 compared to infants with raw scores 90-99 or above 99. Thus, we chose to conduct reliability and validity analyses of Premie-Neuro classifications using a total raw score of ≥ 90 to classify an assessment as normal.

Reliability

Of the original 34 infants, 4 infants were not administered all three subtests of the Premie-Neuro during test-retest reliability testing, and 4 infants were discharged

from the NICU before the test could be administered on day 2. Thus, 26 infants were included in our analyses of short-term test-retest reliability. Using a total raw score of 90 as a cut-point, test-retest reliability of Premie-Neuro classifications was moderate according to criteria described by Landis and Koch (88), and was statistically significant ($Kappa=0.469$, $p=0.005$).

Eight infants were discharged from the NICU before they were given a second weekly Premie-Neuro test. Thus, week-to-week stability of Premie-Neuro classifications was analyzed for 26 infants. Using a total raw score of 90 as a cut-point, 18 infants tested most frequently in the normal classification, 6 infants tested most frequently as questionable, and 2 infants had an equal number of tests in the normal and questionable classifications. The mean percentage of assessments that each infant tested in his or her most frequent classification was 84.5 ($\pm 18.5\%$). Test-retest reliability and stability data with comparisons to data presented in Chapter 2 (analyzed using a raw score cut-point of 100) are summarized in Table 4.3.

Construct validity

Ten infants were categorized as high risk (HR) using the NBRS, while 24 were classified as low risk (LR). Using a total raw score of 90 as a cut-point, a significantly larger proportion of HR infants than LR infants were classified as questionable at the discharge Premie-Neuro assessment ($p=0.000$). Oxygen-dependence at 36 weeks was a significant predictor of Premie-Neuro classifications

($p=0.034$), but gestational age at birth and presence of CNS injury were not significant predictors (Table 4.4).

Predictive validity

One infant's family could not be contacted for term follow-up. Thus, 33 infants were included in analyses for predictive validity of discharge Premie-Neuro classifications at term adjusted age. Using a total raw score cut-point of <90 to classify an assessment as not normal, the discharge Premie-Neuro classifications were significantly associated with NeoNeuro classifications at term ($p=0.045$) and discharge Premie-Neuro classifications were 44.4% sensitive and 92.3 % specific in detecting a not normal NeoNeuro classification at term.

At 3 months adjusted age, five infants' families could not be contacted, three infants' families had relocated, and one infant was deceased. The 25 remaining infants were included in analyses for predictive validity of discharge Premie-Neuro classifications at 3 months adjusted age. Using a total raw score of 90 as a cut-point, discharge Premie-Neuro classifications were significantly association with AIMS ($p=0.035$) but not Infanib ($p=0.130$) classifications at 3 months adjusted age. Discharge Premie-Neuro classifications were 35.3% sensitive and 100% specific in detecting a not normal classifications on the Infanib and 75% sensitive and 85% specific in detecting abnormal classifications on the AIMS at 3 months adjusted age. Table 4.5 describes predictive validity data using a raw score cut-point of <90 to classify an assessment as not normal compared to data presented in Chapter 3, where

predictive validity was calculated using the author-recommended raw-score cut-point of <100.

4.5 Discussion

Although there is no question that preterm infants are at greater risk for developmental delay and disability than their term peers, the majority of preterm babies grow up with no significant developmental disabilities or delays. A disproportionate amount of early intervention dollars are spent to provide services such as physical and occupational therapy to preterm infants (64), but research shows that these services do not result in improved motor outcomes when they are provided to preterm infants simply because they are preterm (8-11). However, when provided to preterm babies who have been identified as high-risk using a standardized assessment tool, studies have shown that these services do have a positive effect on motor outcomes (12, 13). In order to ensure that early intervention services are provided to preterm children in an appropriate, effective, and cost-efficient manner, it is critically important that neonatal care providers have access to standardized tools to objectively identify babies most at-risk for developmental disability and delay and who will, therefore, benefit from early intervention services.

The Premie-Neuro holds promise as a tool that may be used by the clinician for repeated standardized clinical neurological assessments in the NICU. It is brief (takes less than 10 minutes to administer), non-invasive and non-stressful for the infant, and requires minimal training for an experienced clinician. Our previous work

shows Premie-Neuro raw scores are reliable and valid for identifying at-risk infants in the NICU. However, currently recommended Premie-Neuro raw score cut-points for classifying infants as normal, questionable, or abnormal are not reliable and valid and must be interpreted cautiously by the clinician. Although the reliability and validity of raw scores make the Premie-Neuro a valuable tool, it is important that the clinician have some sort of criteria to use to interpret scores, so that they can determine which babies are most at-risk for delays and/or disabilities. A reliable and valid method for classifying Premie-Neuro assessments as normal or not normal—and thus identifying preterm babies as low- or high-risk—would greatly enhance the clinical utility of the tool.

In the present study, our data indicated that prediction of term and 3-month outcomes may not be possible until 33-34 weeks PMA using the Premie-Neuro. Before that time, babies in our study who would later show neurological abnormalities often tested in the normal range, with declining scores only as they approached term adjusted age. This is consistent with our previous work, presented in Chapters 2 and 3, that showed that discharge Premie-Neuro assessments were more reliable and valid than earlier assessments. This does not necessarily mean that earlier Premie-Neuro assessments are not valuable—babies with low early Premie-Neuro scores should certainly be monitored closely and/or provided with services to address deficits—but that an early normal test probably does not effectively rule out neurodevelopmental delay or disability. It is important to note that, in our sample, mean total scores for babies with abnormal follow-up dropped below 90 after 34

weeks PMA and remained <90 through term adjusted age. Thus, a raw score of <90 , particularly at 35-37 weeks PMA, may be indicative of persistent neurological abnormality. Repeated testing through at least 34 weeks PMA is recommended to help ensure that babies with developing abnormal neurological signs are identified, and continued testing through the first months of life using the Premie-Neuro and other standardized tools may determine whether such abnormalities are transient or persistent.

Our study also showed that infants with discharge Premie-Neuro raw scores <90 performed significantly worse on assessments administered at term and 3 months adjusted age than babies with higher scores, particularly compared to babies with scores ≥ 100 . Using a total raw-score cut-point of <90 , rather than using the currently recommended cut-point of <100 , to classify an assessment as not normal resulted in better reliability, validity, and specificity of discharge Premie-Neuro classifications. While further testing with larger samples are needed, at this time it seems most appropriate to suggest that the clinician refer a preterm infant for early intervention services and/or careful follow-up if his or her Premie-Neuro score is <90 at 35-37 weeks PMA. However, it is important to note that—although specificity of discharge Premie-Neuro classifications was better using the raw score cut-point of <90 , sensitivity was worse. For the clinician, this means that using the modified raw-score cut-point of <90 will result in fewer false-positive assessment results, effectively doing a better job at identifying babies who are truly at-risk, but will result more

false-negatives assessments, meaning that a handful of at-risk babies will be missed using the new cut-point.

While our data showed strong prediction for abnormal term or 3-month outcomes in babies with discharge Premie-Neuro scores <90 and normal term and 3-month outcomes in babies with discharge Premie-Neuro scores ≥ 100 , outcomes for babies with discharge Premie-Neuro scores in the 90-99 range were less clear. The majority of babies with discharge Premie-Neuro scores in the 90-99 range will probably develop normally, but there are a handful of those babies for whom neurological delays and/or disabilities may emerge over time. Based on our data, babies with scores in this range should be classified as questionable or suspect and should be re-tested and/or followed carefully to determine whether their borderline scores are indicative of true neurological disability or are simply evidence of mild delays and/or transient, resolving neurological signs. Premie-Neuro assessment results for these babies may be combined with other data—such as results of neuroimaging—and clinical judgment for making decisions on whether to refer these babies for early intervention services.

A limitation of this study is the relatively low sample size. We tested 34 infants in this study and, due to circumstances such as early NICU discharge or inability to contact families at follow-up, some analyses included as few as 26 subjects. Perhaps because of this, there was little variation in Premie-Neuro raw score between infants, and no infant scored in the abnormal range (raw score <70 as recommended by the Premie-Neuro authors) during the course of the study.

Currently, the Premie-Neuro recommends 3 classifications based on raw-score cut-points (normal: ≥ 100 ; questionable: 70-99; abnormal: < 70). Because the infants in our study only tested in two classifications—questionable and normal—we collapsed all assessments into 2 classifications (normal and not normal) for our analyses. Future studies with larger groups of infants may result in greater variation in performance, allowing further analyses of Premie-Neuro classifications to be done using a 3-classification system rather than 2.

Another limitation of this study is that the participants were only followed through 3 months adjusted age. Over 25% of the infants in our study had abnormal assessment results at term and/or 3 months adjusted age. Research shows that major disability, such as cerebral palsy, generally occurs in 7-20% of preterm infants, with the average generally falling around 10-15% (3, 32, 41-44). Thus, it is unlikely that all 9 of the infants in our abnormal follow-up group will truly have significant neurodevelopmental abnormalities. Follow-up testing of the infants in this study will continue through at least 2 years of age to determine if abnormalities tend to resolve in some babies and, if so, if the Premie-Neuro was able to predict which babies would have true, persistent neurodevelopmental deficits.

4.6 Conclusions

The results of this study show that, after approximately 34 weeks PMA, the Premie-Neuro does differentiate between infants with normal versus abnormal performance on standardized assessments administered at term and 3 months adjusted

age. Infants with discharge Premie-Neuro scores <90 performed worse on all assessments administered at term and 3 months adjusted age than infants with scores ≥ 100 , but there were not clear differences among infants with scores in the 90-99 range, particularly at 3 months adjusted age. Using the Premie-Neuro raw score cut-point of <90 to classify an assessment as not normal resulted in better reliability, validity, and specificity versus using the currently recommended cut-point of ≥ 100 . Based on these data, we recommend that infants with Premie-Neuro raw scores <90 at 35-37 weeks PMA be classified as not normal or at-risk and be referred for more in-depth, specialized testing and/or early intervention services. Babies with raw scores in the 90-99 range should be classified as questionable, and babies with raw scores ≥ 100 be classified as normal. As the Premie-Neuro becomes more widely used, future research will be needed to determine if these cut-points are indeed appropriate for identifying at-risk infants and to examine whether the Premie-Neuro may be modified to enhance reliability and validity at earlier ages in the NICU.

	Sex	GA at birth, weeks	BW, grams	Birth order	Race/ethnicity
Normal at follow-up (n=24)	17 male 7 female	30.0 (3.6)	1544.6 (695.3)	11 singletons 10 twins 3 triplets	19 White 1 African-American 2 Asian 2 Hispanic
Abnormal at follow-up (n=9)	8 male 1 female	26.1 (1.8)	783.7 (265.2)	4 singletons 4 twins 1 triplet	8 White 0 African-American 0 Asian 1 Hispanic
All participants (n=34)	25 male 9 female	29 (3.7)	1343.2 (696.3)	16 singletons 14 twins 4 triplets	27 White 2 African-American 2 Asian 3 Hispanic

Table 4.1. Description of study participants. Data are mean (standard deviation). One infant could not be contacted for term or 3-month follow-up testing, and was therefore not included in either the normal or abnormal follow-up group. GA=Gestation Age; BW=Birth Weight.

	PMA (weeks)						
	31	32	33	34	35	36	37
Abnormal at follow-up (n=9)	101.5 (10)	101 (3.8)	95.6 (6.7)	91.3 (6.2)	87.7 (3.9)	80 (6.3)	83 (7.2)
Normal at follow-up (n=24)	101.7 (10)	98.9 (3.8)	99.7 (5.6)	97.3 (5.1)	98.8 (6.2)	96.3 (8.3)	99.3 (9.4)
All participants (n=33)	101.6 (9.5)	99.6 (3.7)	98.8 (6.2)	94.9 (6.1)	95.6 (7.6)	90.6 (10.9)	93.5 (11.7)

Table 4.2. Premie-Neuro raw scores for infants with abnormal follow-up

assessments (abnormal NeoNeuro at term or abnormal Infanib or AIMS at 3 months AA), normal follow-up assessments (normal NeoNeuro at term and normal Infanib and AIMS at 3 months adjusted age), and all participants at each week PMA. Data are mean (standard deviation). Data from infants assessed at PMA weeks 29 and 30 were omitted as there was only one data point at 29 weeks and no data points for 30 weeks in the abnormal follow-up group. PMA=post-menstrual age, AA=adjusted age.

	Total Raw Score Cut-point	
	90	100*
One-day test-retest reliability (n=26)	$\kappa=0.469$	$\kappa=0.137$
Stability (% weekly assessments in most frequent classification, n=26)	84.5%	79.5%

Table 4.3. Reliability and stability of Premie-Neuro classifications using raw

score cut-points of 90 and 100. Raw scores at or above the cut-point were classified as normal, while raw scores below the cut-point were classified as not normal. Using the raw score cut-point of 90 resulted in improved test-retest reliability and stability of the Premie-Neuro.

***Data from Chapter 2.**

	Total Raw Score Cut-point	
	90	100*
Proportion of HR vs. LR infants classified as “questionable” (Fisher’s exact test)	p=0.000**	p=0.191
Significant predictors of Premie-Neuro classifications (Logistic regression with CNS abnormalities, GA at birth, and oxygen-dependence at 36 weeks entered as independent variables)	Oxygen-dependence at 36 weeks (p=0.034**)	Oxygen-dependence at 36 weeks (p=0.048**)

Table 4.4. Construct validity of Premie-Neuro classifications using raw score

cut-points of 90 and 100. Raw scores at or above the cut-point were classified as normal, while raw scores below the cut-point were classified as not normal.

There was a significantly larger proportion of high-risk infants classified as not normal using a cut-point, but no group differences using a cut-point of 100.

Oxygen-dependence at 36 weeks, but not CNS abnormality, was predictive of Premie-Neuro classifications using both cut-points. GA at birth was not included in either regression model due to lack of significance. HR=High Risk Group; LR=Low Risk Group; CNS=Central Nervous System; GA=Gestational Age.

***Data from Chapter 2.**

****p<0.05**

	Total Raw Score Cut-point	
	90	100*
NeoNeuro (term)		
Proportion of infants classified as not normal on PN who were also classified as not normal on the NN (Fisher's exact test)	p=0.045**	p=0.008**
Sensitivity	53%	85%
Specificity	92%	69%
Infanib (3 months adjusted age)		
Proportion of infants classified as not normal on PN who were also classified as not normal on the INF (Fisher's exact test)	p=0.130	p=0.371
Sensitivity	41%	61%
Specificity	100%	71%
Alberta Infant Motor Scale (3 months adjusted age)		
Proportion of infants classified as not normal on PN who were also classified as not normal on the AIMS (Fisher's exact test)	p=0.035**	p=0.093
Sensitivity	75%	100%
Specificity	80%	60%

Table 4.5. Predictive validity of the Premie-Neuro using raw score cut-points of 90 and 100. Raw scores at or above the cut-point were classified as normal, while raw scores below the cut-point were classified as not normal. Using a raw score cut-point of 90 improved predictive validity and specificity of the Premie-Neuro at term and 3 months adjusted age, but resulted in worse sensitivity.

PN=Premie-Neuro; NN=NeoNeuro; INF=Infanib; AIMS=Alberta Infant Motor Scale.

***Data from Chapter 3, **p<0.05**

4.7 Figure Legend

Figure 4.1. Mean Premie-Neuro total raw scores from 29-37 weeks post-menstrual age (PMA) in infants with normal term and 3-month follow-up assessments versus infants with abnormal results on assessments administered at term or 3 months adjusted age. After approximately 34 weeks PMA, Premie-Neuro scores for infants with normal follow-up remained stable, while scores for infants with abnormal follow-up worsened.

Figure 4.2. Premie-Neuro total raw scores from 29-37 weeks post-menstrual age (PMA) for individual subjects with (a) normal term and 3-month follow-up assessments and (b) abnormal results on assessments administered at term or 3 month adjusted age.

Figure 4.3. Mean performance on (a) the NeoNeuro at term and (b) the Infanib and (c) the Alberta Infant Motor Scale (AIMS) at 3 months adjusted age for infants with discharge Premie-Neuro scores <90, 90-99, and >99. Infants with discharge Premie-Neuro raw scores <90 performed significantly worse than infants with discharge Premie-Neuro scores >99 on all term and 3-month follow-up tests, and performed significantly worse than infants with discharge Premie-Neuro scores 90-99 on the NeoNeuro at term and the Infanib at 3 months adjusted age. There was no difference between infants with discharge Premie-Neuro scores 90-99 and infants with discharge Premie-Neuro scores >99 at term or 3 months adjusted age. * $p<0.05$

Figure 4.1

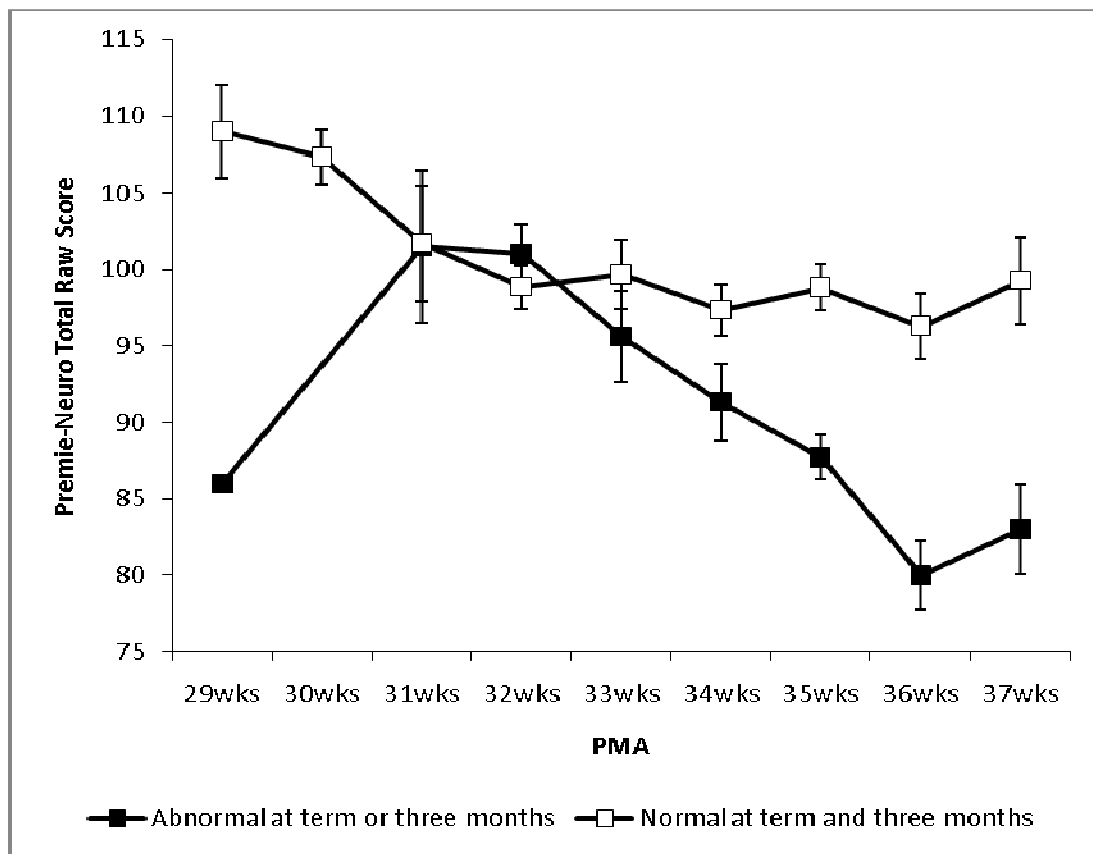
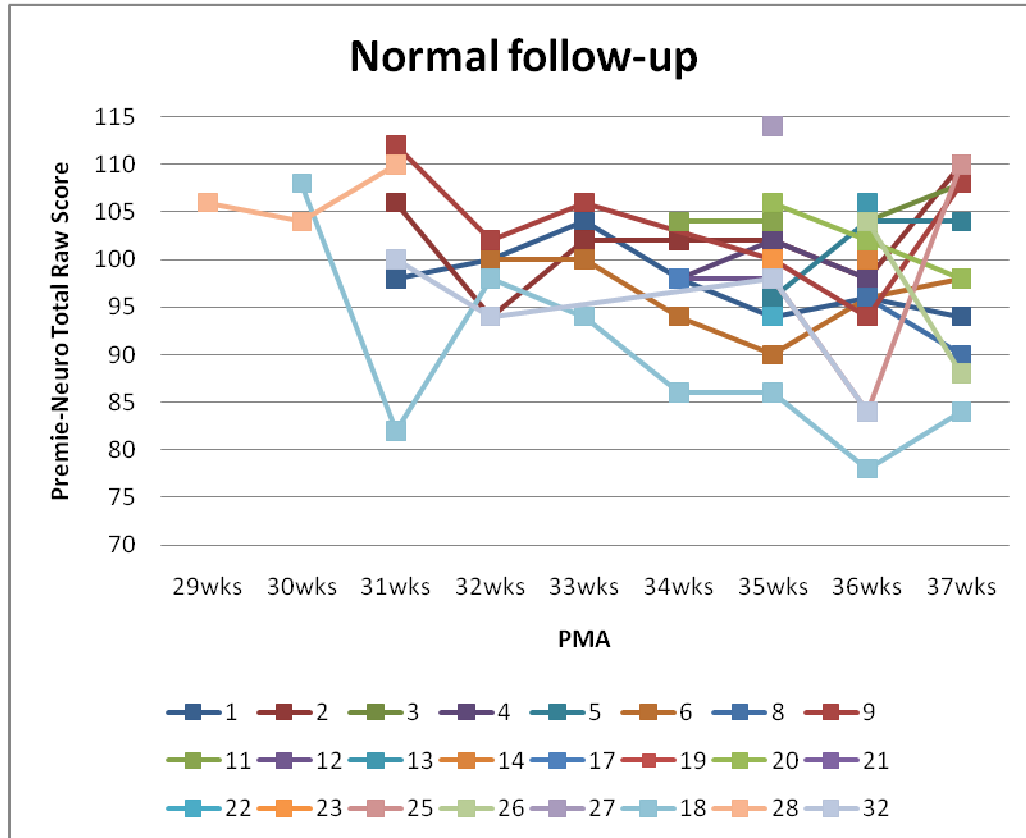


Figure 4.2

a.



b.

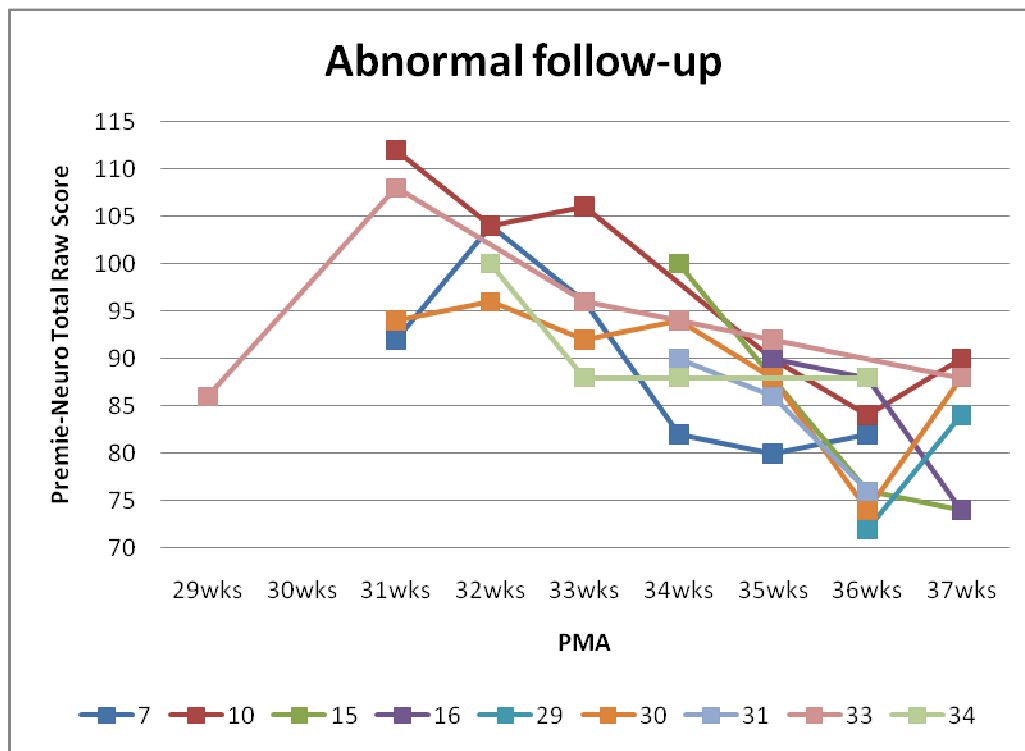
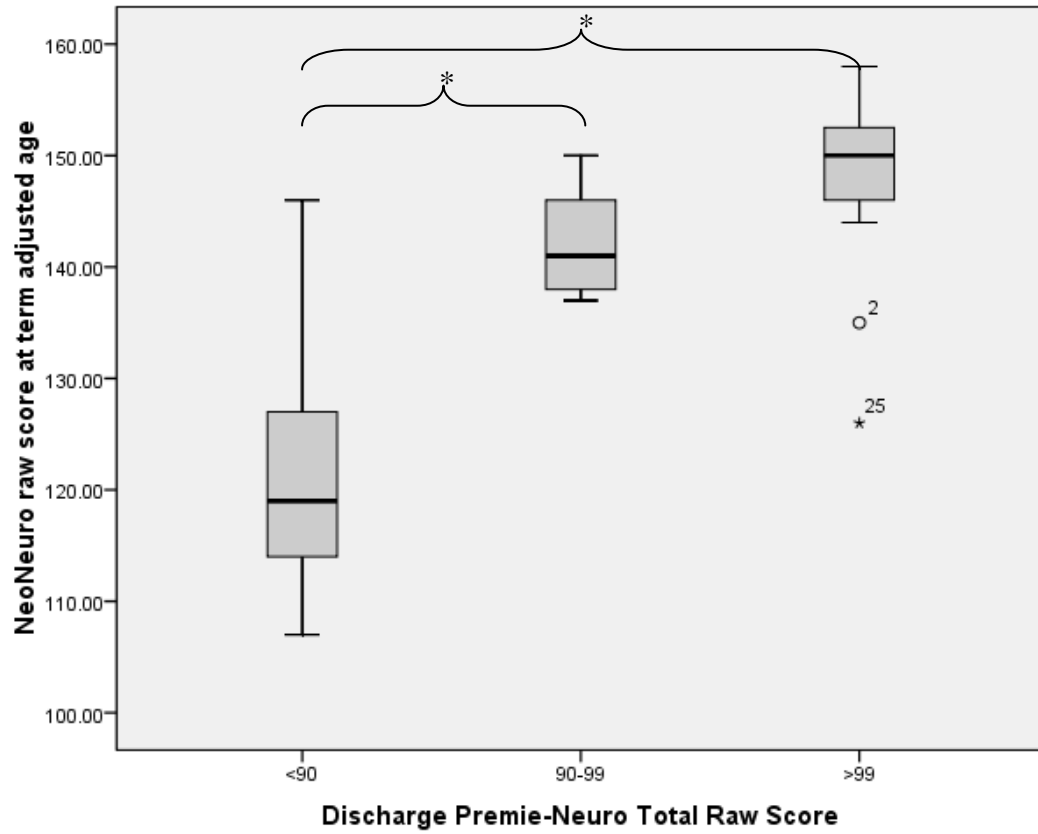
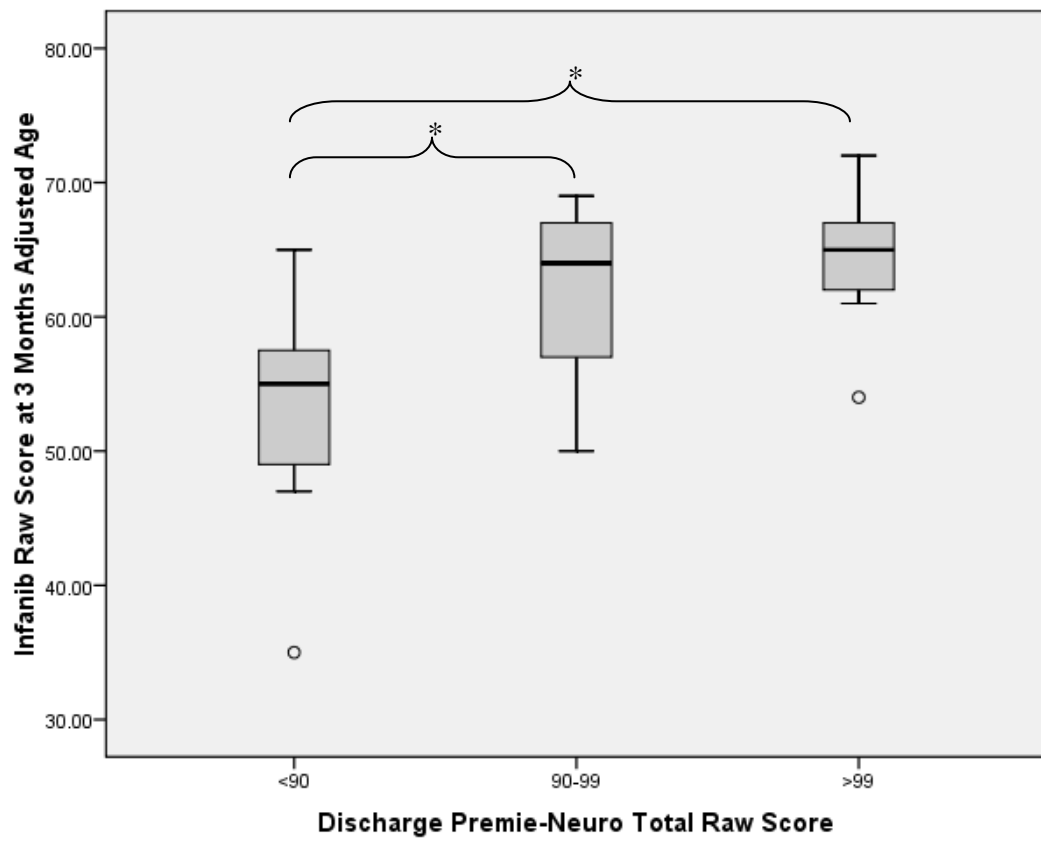


Figure 4.3

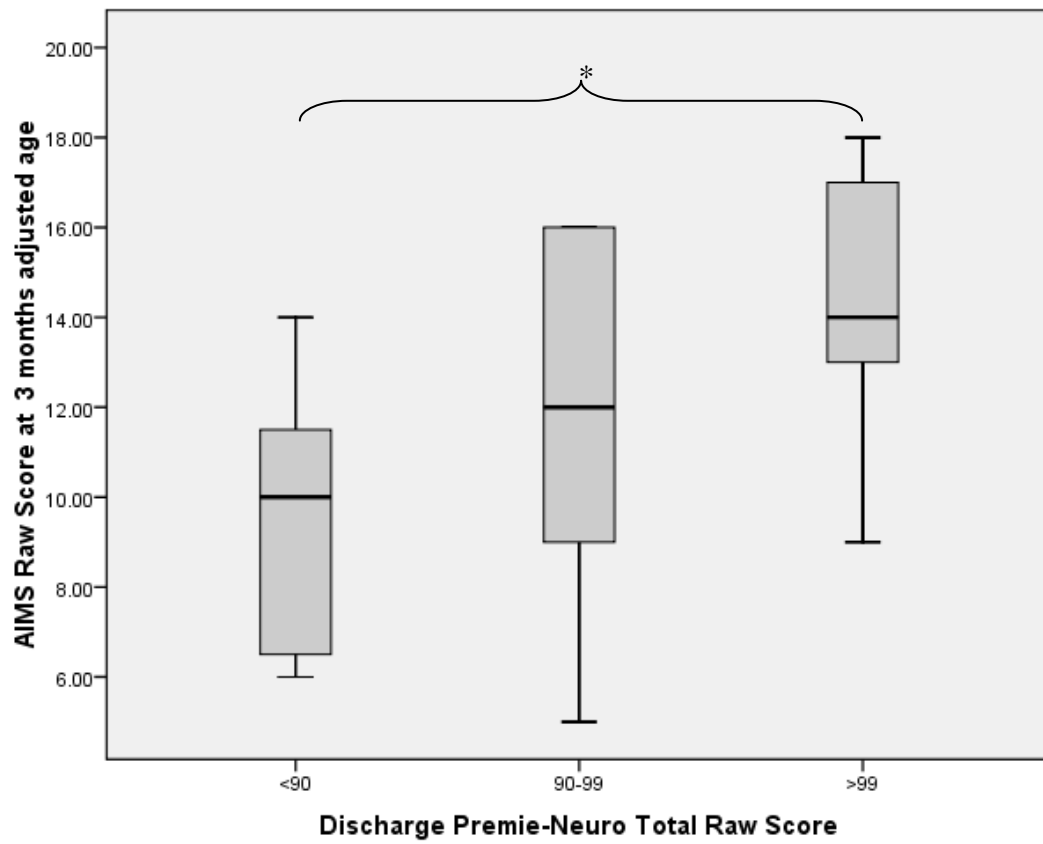
a.



b.



c.



Chapter 5

Summary and Conclusions

5.1 Summary of findings

The primary purpose of the work presented in this dissertation was to determine the reliability and validity of the Premie-Neuro, a relatively new clinical test of neurological development for the preterm infant age 23-37 weeks post-menstrual age (PMA). A secondary goal was to determine what modifications of the test, if any, may be recommended to potentially improve the reliability and validity of the tool. Prior to the present study, internal consistency of the Premie-Neuro had been established (14) and pilot work had shown that the Premie-Neuro may be predictive of outcomes through 6-8 months adjusted age (79), but no other study had shown whether the Premie-Neuro had acceptable reliability and validity for clinical use. Overall, the body of work presented in this dissertation showed that Premie-Neuro raw scores had acceptable interrater and test retest reliability (Chapter 2), good construct validity (Chapter 2), and the ability to predict outcomes at term and 3 months adjusted age (Chapter 3). Although Premie-Neuro classifications (normal, abnormal, and questionable) based on currently recommended raw score cut-points were not as reliable and valid as raw scores, we used our data to recommend modifications to Premie-Neuro classifications that seemed to improve their reliability and validity (Chapter 4). These results may allow for earlier identification of neurodevelopmental disabilities in preterm infants in the neonatal intensive care unit (NICU), and better decision-making for referrals for specialized early intervention services, such as physical or occupational therapy, in the NICU and/or at NICU discharge.

Chapter 2. The Premie-Neuro: A reliable and valid clinical neurological test for preterm infants in the NICU.

The purpose of Chapter 2 was to determine if Premie-Neuro raw scores and classifications (normal, abnormal, or questionable) based on raw score cut-points recommended by the assessment's authors had acceptable reliability and validity for clinical use. We found that interrater and test-retest reliability of raw scores—although lower than reliability previously reported for other assessments designed for older preterm and term newborns (65, 68, 72, 90, 92, 95, 96)—was fair to moderate, representing clinically significant reliability of the Premie-Neuro for this population. Construct validity of the Premie-Neuro raw scores was strong as the assessment was able to differentiate between infants at different risk for poor outcomes. Although stability of Premie-Neuro classifications was good from week to week, interrater and test-retest reliabilities and construct validity of Premie-Neuro classifications were inadequate. Thus, our work in Chapter 2 suggested that, as the Premie-Neuro is currently structured, raw scores may be used to reliably identify at-risk babies in the NICU, but classifications must be interpreted carefully.

Chapter 3. Predictive Validity of the Premie-Neuro at Term and 3 Months Adjusted Age.

One of the biggest challenges in neonatal neurological assessment is determining which babies are demonstrating true, persistent neurological abnormalities and which babies' abnormal neurological signs will resolve over time.

Thus, predictive validity—the ability for performance on an assessment to predict future performance on a different assessment—is of considerable importance to the clinician. The goal of Chapter 3 was to determine whether Premie-Neuro performance in the NICU could predict performance on the NeoNeuro at term adjusted age and the Infant and Alberta Infant Motor Scale (AIMS) at 3 months adjusted age. We found that Premie-Neuro classifications and early Premie-Neuro raw scores generally had poor predictive value. However, Premie-Neuro raw scores obtained shortly before discharge from the NICU (generally after 34-35 weeks PMA) were significantly associated with outcomes at term and 3-months adjusted age. This provided strong evidence that, as infants approach term adjusted age, the Premie-Neuro raw scores were able to identify neurological abnormalities that would persist through at least 3 months adjusted age. Thus, discharge Premie-Neuro raw scores may be used to make recommendations for follow-up and/or necessity of early intervention services.

Chapter 4. Modifying raw score cut-points improves reliability, validity, and predictive value of Premie-Neuro classifications.

Despite the fact that the Premie-Neuro raw scores had acceptable reliability and strong construct and predictive validity, reliability and validity of the Premie-Neuro classifications were generally weak (Chapter 2, Chapter 3). Thus, in Chapter 4, we retrospectively analyzed the data presented in Chapters 2 and 3 to determine if the raw score cut-points used for classifying a Premie-Neuro assessment as normal,

questionable, or abnormal could be modified to improve reliability and validity, thereby enhancing the clinical usefulness of the tool. We found that infants with Premie-Neuro raw scores <90 after approximately 34 weeks adjusted age performed significantly worse on all assessments at term and 3 months adjusted age than infants with Premie-Neuro raw scores ≥ 100 and significantly worse on the NeoNeuro at term and the Infanib at 3 months adjusted age than infants with scores in the 90-99 range. Thus, we re-analyzed all data using a raw score cut-point of <90 to classify an assessment as not normal and found improved reliability, validity, and specificity using the new cut-point. Based on our data, we recommended that the clinician may consider an infant not normal or at-risk if he or she scores <90 on Premie-Neuro assessments after 34 weeks PMA, and that the infant should be classified as questionable and re-tested or closely monitored if he or she scores 90-99. Infants with raw scores >100 may be classified as normal. However, we caution that more research needs to be done to evaluate the reliability and validity of the Premie-Neuro using the new cut-points. Further, the clinician should note that sensitivity of Premie-Neuro classifications was worse using the raw score cut-point of <90 to classify an assessment as abnormal. Thus, interpreting Premie-Neuro assessments using the raw score cut-point of <90 increases the risk of missing an infant who will later demonstrate neurological abnormalities.

5.2 Clinical implications

The preterm birth rate has risen 20% since 1990. In 2006, over 540,000 infants were born preterm in the United States (1). Many preterm infants go on to develop normally. However, several recent reviews report that up to 25% of preterm infants are eventually diagnosed with major disabilities such as cerebral palsy, and the majority of preterm infants may exhibit milder dysfunction in areas such as motor coordination, behavior, reading and verbal skills, and cognition (3, 21, 44). In the past, neonatal and pediatric care providers have had very few tools at their disposal for early identification of high-risk infants, and it has been long thought that neurological disability cannot be diagnosed until at least 1-2 years of age. However, there is mounting evidence that standardized testing in the NICU can predict neurodevelopmental outcomes in preterm babies (57, 58). In this body of work, we have shown that the Premie-Neuro is a reliable and valid clinical neurological test and that it is predictive of outcomes through 3 months adjusted age. Because it can be administered to infants as young as 23 weeks post-menstrual age in less than 10 minutes without removing the infant from his or her bed and without interfering with mechanical ventilation or electronic monitoring, it is ideal for repeated use during an infant's NICU stay to track his or her neurological development and discriminate between transient or persistent neurological signs. The results of our work have implications for the way preterm infants are assessed and cared for in the NICU and followed-up after discharge.

How early can we identify developmental disability and delay in the preterm infant?

In the present study, we found that the Premie-Neuro did not discriminate between high- and low-risk infants (Chapter 2), nor did it discriminate between infants who would later have normal versus abnormal follow-up outcomes (Chapter 4), until approximately 34-35 weeks post-menstrual age. Before that time, mean raw scores for all groups of infants tended to be in the normal or near-normal range. At 33-34 weeks PMA, mean raw scores for low-risk infants and those with normal follow-up at term and 3 months adjusted age remained steady at approximately 95-100, while means scores for high-risk infants and those with abnormal follow-up at term or 3 months adjusted age fell to approximately 80-85 as they approached 37 weeks PMA. We do not believe that this means that early Premie-Neuro testing is not valuable or should not be done. However, it does mean that the Premie-Neuro may not be sensitive enough to detect abnormalities in some very young preterm infants, and an early normal result should not be used to rule out neurological abnormality. Repeated testing should be done through at least 34 weeks PMA to ensure that abnormal neurological signs do not emerge. However, an early abnormal result should not be ignored and may indicate that further, more in-depth testing or closer follow-up may be indicated.

Striking a balance between sensitivity and specificity

Preterm infants are a highly variable population, both within and between babies. When an infant demonstrates a neurological abnormality, it is difficult to

know if that abnormality will persist and lead to significant neurological dysfunction or if it is a transient neurological sign that will resolve (2, 3, 21). Neonatal neurological assessment tools are often not sensitive enough to distinguish between persistent and transient neurological signs, and although these assessments generally have good negative predictive value (children who test in the normal range do well), they are notorious for producing high numbers of “false positives” (65). That is, infants are often identified as having a neurological abnormality when they will in fact go on to be neurologically normal. In this study, there was a tendency to over-identify assessments as abnormal using the author’s recommended raw score cut-point of 100. However, when we lowered the raw score cut-point to 90, we found the reverse was true. There was better specificity and therefore fewer false-positives, but worse sensitivity and more false-negatives (the test tended to miss infants who would later score in the not normal range on testing at term and 3 months adjusted ages).

When attempting to identify infants with neurological abnormalities, there may be benefits associated with a test with good negative predictive value and a tendency toward false positives. It allows reassurance when an infant is tested as neurologically normal, and causes examiners to exercise caution and closely monitor “borderline” children whose abnormalities may resolve. One may argue that it is better to provide unnecessary early services than to deny services to a child who needs them, and that is often the position of NICU clinicians and early interventionists. However, research does not provide strong support for specialized early intervention services for broad groups of preterm infants (10, 99), and false-

positives can place undue emotional stress on an infants' family. In a time where state and federal dollars are scarce, clinicians may need to shift their thinking away from providing early intervention services to preterm infants "just in case" and toward efficiency and proven effectiveness by using more specific standardized assessment tools to better identify high-risk infants who will truly benefit from services. Using a Premie-Neuro raw score of <90 to classify an infant as high-risk at or near NICU discharge, as proposed in Chapter 4, resulted in higher specificity and fewer false-positive results, making the tool unique compared to other neonatal neurological assessments that tend toward higher sensitivity. We believe that the Premie-Neuro and this more conservative approach to interpreting raw scores should be used to make better clinical decisions about services for preterm babies. By designating infants with scores of 90-100 as a neurologically questionable group that should be re-tested and closely monitored, we hope to minimize the number of false negative tests by capturing babies with borderline neurological function and whose neurological abnormalities may emerge over time.

Differentiating between transient and persistent neurological abnormalities

As mentioned previously, serial neurological assessment of preterm infants is critical for distinguishing between persistent and transient abnormalities. However, many currently available assessments are not feasible for repeated clinical use due to time constraints and limitations in the ages at which the tests may be used. Because the Premie-Neuro may be administered in less than 10 minutes to all infants in the

NICU (from the limits of viability to term adjusted age), it is ideal for providing repeated assessments throughout an infant's NICU stay. This allows the neonatal clinician the unique opportunity to track neurological development and differentiate between persistent and resolving neurological abnormalities.

Using Premie-Neuro test results to make decisions about supports, services, and follow-up

Standardized neurological assessments for preterm and newborn infants traditionally classifications of normal, abnormal, and a third category—usually called transient or questionable—for infants who score somewhere in between. While it is important to recognize that there are significant numbers of infants with neurological signs that are not necessarily “abnormal,” it is sometimes difficult for the clinician to interpret these assessments for clinical decision-making, particularly when infants test in the “questionable” or “transient” classification. From a clinical perspective we felt that it was important to determine what Premie-Neuro classifications meant in terms of long-term outcomes. We found that, after approximately 34 weeks PMA, infants with discharge Premie-Neuro raw scores <90 had significantly worse outcomes at term and 3 months adjusted age than scores ≥ 100 . However outcomes for infants with scores 90-99 were less clear. Using these data, we proposed keeping the traditional 3-classification system, classifying scores <90 as abnormal and scores ≥ 100 normal, but we recommended narrowing the “questionable” range to include scores from 90-99. For the clinician, that means that infants who score repeatedly

<90 should be considered “high-risk” and referred for specialized early intervention services or for further, more in-depth testing. Infants with scores >99 after 34-35 weeks PMA may be “low risk” with no need for extra services and supports, but should still be provided routine follow-up care to make sure that no neurological abnormalities emerge. Our longitudinal data showed that, with multiple assessments, many infants with early Premie-Neuro raw scores 90-99 had a tendency to remain steady or rise as they approached term if they later had normal outcomes, but tended to drop below 90 if their outcomes at term and/or 3 months adjusted age were abnormal. Thus, we recommend that infants with discharge Premie-Neuro scores 90-99 should be monitored closely to determine neurodevelopmental progress. In all instances, but particularly for infants who test repeatedly in the 90-99 range, it is important to combine results of standardized assessments with clinical and medical data to make decisions about services and follow-up.

Using the Premie-Neuro to evaluate the effectiveness of neonatal care and interventions

In addition to allowing the clinician to make better decisions about recommending special supports and services for preterm infants, we believe the Premie-Neuro may be used to evaluate the effectiveness of those interventions. Although clinicians are able to use imaging techniques such as cranial ultrasound or MRI to monitor the *structure* of the preterm infant brain, the Premie-Neuro provides a unique opportunity to measure the *function* of the preterm infant brain over time.

This has important implications for both clinical practice and research. Because it is a reliable and valid measure, the Premie-Neuro may be used to quantify brain function prior to initiating common NICU intervention such as music therapy, kangaroo care, infant massage, or a positioning program. Repeated Premie-Neuro assessments during the course of these interventions or after the intervention period has concluded would provide information on the developing brain's response to these interventions, giving some insight as to their effectiveness. Furthermore, the Premie-Neuro could be used to assess the effect of developmental care—promoting physiological flexion and controlling NICU sounds, light, and temperature to make the environment more womb-like—on the preterm infant's developing brain. For the clinician, the Premie-Neuro would provide a baseline against which to measure an individual infant's change in brain function over time in response to an intervention. For the researcher, the Premie-Neuro could be used in randomized controlled trials to determine the effect of NICU interventions on groups of infants.

5.3 Limitations

Setting of the study

All Premie-Neuro testing was conducted in the NICU at bedside rather than in a standardized lab, clinic, or exam room. Because the infant was not removed from his or her crib or isolette or the NICU, testing time was often flexible and the infant was disturbed very little by the testing. However, this did not allow for control of environmental factors such as sound, light, and temperature; caregiver presence;

wires and tubes (electronic monitoring, oxygen, and/or ventilation) in the testing area; or height, incline, or type of testing surface. Follow-up testing was conducted in the NICU, NICU follow-up clinic, or infant home, depending on caregiver preferences. Again, this did allow for flexibility of testing time but did not allow for control of the testing environment. Testing the infants in a controlled environment may have strengthened reliability and, to a lesser extent, validity data. However, we believe that our study provided us with the unique opportunity to observe and test infants in their natural environments, and that these data represents reliability and validity that would be obtained during testing in a “real-life” situation.

Because of the fragile nature of the infants in the NICU environment and their unique medical needs, it was not possible to control for medications or medical/developmental care provided to the participants in this study. Further, because the infants’ mothers were not part of the study, access to prenatal history data was limited. Thus, we were unable to control for factors that may have influenced prenatal development such as poor prenatal nutrition or prenatal drug exposure.

Subjects and subject characteristics

Thirty-four infants participated in this study, 4 more than we initially proposed. This was a sample of convenience recruited from the NICU at Overland Park Regional Medical Center. Unfortunately, our sample was relatively homogenous and consisted primarily of white (79.4%) and male (73.5%) babies. Although the racial/ethnic background of the subjects recruited for our study is

probably reflective of the community in which the subjects were recruited (Johnson County, Kansas), we are unsure why our sample was so heavily male. There is conflicting evidence of whether race or sex play a significant role in neurodevelopment (24, 37, 59, 62, 74), and future studies are needed to determine if there is a race or gender bias associated with the Premie-Neuro.

Although 16 of the 34 participants in this study were born prior to 28 weeks completed gestation, no testing on any infant was conducted until at least 29 weeks PMA. This was largely due to health concerns and the difficulty obtaining parental consent for testing very young, medically fragile infants. Although the Premie-Neuro is designed for infants as young as 23 weeks PMA, the results of this study cannot be generalized to infants less than 29 weeks PMA. However, the difficulty testing very young preterm infants in this study is probably reflective of the difficulty testing very young, medically fragile infants in any clinical setting.

We anticipated an attrition rate of up to 10% during the course of this study (100). However, NICU discharge during Premie-Neuro testing resulted in many analyses including as few as 22 subjects. Only one infant (2.9%) was lost to follow-up at term adjusted age, but 9 (26.5%) were lost at 3 months adjusted age. Although the attrition rate was higher than we anticipated, this was mediated somewhat by our larger-than-proposed sample size, and we believe that our sample size provided adequate power for the analyses conducted in the present study. Again, this loss of babies to follow-up may be reflective of the challenges faced by neonatal clinicians in their follow-up of preterm children.

Examiner training and instruction

According to *The Premie-Neuro Examination Instruction Manual*, “The evaluator should study the items in the protocol and the scoring sheet prior to administering the evaluation. It should be practiced repeatedly until the examiner feels comfortable with the assessment. For data collection, it is recommended that training on the examination occur and the examiners obtain an interrater reliability of 0.90 for scoring of the items before data collection occurs. A video CD or DVD of the Premie-Neuro examination is available for viewing (78).” In the present study, the examiner and the observer studied the items and the scoring sheet and practiced the examination independently prior to beginning the study. However, training to the 0.90 interrater reliability criteria was not done so that we could estimate the true, clinical interrater reliability of the Premie-Neuro. On our sample of 15 infants who were tested for interrater reliability, the intraclass correlation coefficients for interrater reliability were 0.556 for the Neurologic and Movement subtests and 0.391 for all 3 subtests (Chapter 2). Item agreement (percentage of items for which the examiner and observer assigned the infant the same score) was 72.1% (Unpublished data from Premie-Neuro study). Although these reliability estimates are clinically acceptable for this fragile population, our data indicate that it may take more than 15 assessments to train to 0.90 interrater reliability or that this level of interrater reliability may be difficult to achieve with this assessment in this population.

In addition to the inherent lack of neurodevelopmental stability in this population, another reason for lower interrater reliability than estimated may be due to the item descriptions included in *The Premie-Neuro Examination Instruction Manual* or the subjective nature of scoring for some of the items. The examiners in this study did not note this as a significant limitation for the Neurologic and Movement subtests, as scoring these items is typically quantitative (i.e. measuring a joint angle or tracking a number of movements per minute). However, the examiners in this study reported that much of the scoring for items in the Responsiveness subtest seems somewhat subjective. For example, when administering the “Posterior Neck” item, the examiner holds the infant in a supported sitting position and lets the head fall forward. If the infant does not lift the head, the examiner must determine if the infant (a) made no attempt to raise head or (b) tried but could not raise the head (78). There is no description of what constitutes an “attempt” or “try,” and it is possible that this could be interpreted differently by different examiners. In our study, where interrater reliability was studied using a hands-on examiner and an observer, it was extremely difficult for the observer to differentiate between “no attempt” and a “try” when there was no movement to observe. This lack of clear scoring criteria for items in the Responsiveness subtest is reflected in our data, as the interrater reliability for two subtests was much higher than for all three subtests (Chapter 2). More specific scoring criteria in the instruction manual and/or the instruction video may improve reliability of the tool.

Test items

In our study, we noted that predictive validity of the Premie-Neuro was much better as infants approached term adjusted age than at earlier testing (Chapters 3 and 4). This is in agreement with research showing little individual stability in motor activity and behavior prior to 34 weeks PMA, but improving stability from 35-44 weeks PMA (101, 102) and work by Korner and colleagues, who identified 32 weeks PMA as a “...turning point at which preterm infants behave more like term infants than fetuses (70).” Based on that work as well as the work presented here, it may be argued that infants younger than 32-34 weeks PMA may need to be assessed using a different type of test than infants older than 34 weeks PMA. The Premie-Neuro addresses this problem to some degree by adjusting the score on each item based on PMA and giving a shorter form of the assessment to infants less than 28 weeks PMA or on a ventilator, but there are a handful of items that may not be sensitive enough to detect abnormalities until an infant is 32-34 weeks PMA. In fact, there are several items, particularly those in the Responsiveness subtest, tended to have this effect up to 32-34 weeks PMA. For example, for the “Anterior Neck” item, the infant is held in a sitting position and the head is allowed to fall backward. The examiner waits for the infant’s response. Until the age of 34 weeks PMA, “no attempt to raise head” is a normal response. Thus, an infant struggling with head control may not be identified by this item until 34 weeks PMA if they are still unable to attempt to raise the head.

Conversely, there are many items included in the Premie-Neuro—primarily those that assess active and passive tone—that penalize the infant for performing too

well. For example, if an infant is administered the “Anterior Neck” item described in the previous paragraph at 33 weeks PMA, it is considered an abnormal response—indicative of excessive tone—if the infant holds the head upright and maintains that position. However, this is considered a normal response for an infant at 36 weeks PMA. Just like a normal 4-year-old child can be more mature than average, a normal 33-week preterm infant can also be more mature than other same-age infants. In the example given, a very mature preterm infant would be given an abnormal score on the “Anterior Neck” item at 33 weeks PMA even if the infant’s performance were indicative of a very mature brain functioning at closer to the level of 36 weeks PMA, and not abnormally high muscle tone. This early penalization of infants who would later score extremely well on the test may be an additional reason why the Premie-Neuro validity improved significantly as infants approached term adjusted age.

5.4 Future directions

In the body of work presented in this dissertation, we established that the basic psychometric properties—interrater reliability, test-retest reliability, construct validity, and predictive validity—of the Premie-Neuro are acceptable for clinical use. However, more work is needed to further determine long-term predictive validity of the tool. There are many unanswered questions about “normal” preterm infant development as well as the relationship motor development and preterm infant physical growth and physiological stability. Little work has been done to determine the effect of specialized care and NICU interventions on brain development and

function. And there are few data on the utilization of early intervention services by NICU graduates, and the effectiveness of such interventions. The following future studies may address these questions.

Compilation of normative data for infants in the NICU

There is a paucity of normative developmental data on preterm infants, particularly those less than 32 weeks PMA. Because we tested infants as young as 29 weeks PMA, the data collected in this study may be used to begin compiling normative data for very preterm and extremely preterm infants. Performance on each individual test item could be tracked according to PMA, comparing infants with abnormal follow-up to infants with normal follow-up. This would give insight into the true range of normal passive and active tone, reflexes, movements, and responsiveness in young preterm infants and may allow for more precision in determining normal versus abnormal performance on individual Premie-Neuro items. Future testing of greater numbers of babies may allow for norm-referencing of the Premie-Neuro so that each infant's score may be assigned a percentile rank to compare his or her performance to that of same-age peers. Data collected on infant weight, length, and head circumference may be used to determine the relationship between physical growth and neurological development. Physiological data (heart rate and oxygen saturation) collected during Premie-Neuro assessments may be used to quantify physiological stress experienced by preterm infants during NICU testing

and to determine the normal range of physiological responses to handling and developmental testing in young preterm babies.

Responsiveness of the Premie-Neuro to specialized early intervention services provided in the NICU

Because preterm infants in the Overland Park Regional Medical Center NICU do not routinely receive early intervention physical therapy while hospitalized, the data collected in this study characterizes neurological development in preterm infants who are receiving routine hospital care and no additional physical therapy interventions. Thus, the infants tested in this study may represent a control group that can be used for comparison in future experiments studying the effects of specialized early intervention therapy in preterm infants, as well as the Premie-Neuro's responsiveness to such interventions.

Long-term predictive validity of the Premie-Neuro

In the present study, infants were followed and tested through 3 months adjusted age. Although this gave us an estimate of the short-term predictive validity of the Premie-Neuro, there are still questions as to whether the Premie-Neuro has long-term predictive value. Recent reviews by Allen, Aylward, and Bracewell and Marlow summarize the general consensus among neonatal care providers that true neurological dysfunction and delay may not be able to be detected in the preterm infant until at least 18-24 months of age (2, 3, 21, 43). This is due to the transient

neurological abnormalities often exhibited by preterm children—particularly in the first year of life—as well as the presence of developmental delays that often resolve in this population by two years of age. Our hope is that the data in this study will help provide some insight into factors that may be predictive of long-term neurodevelopmental deficits at young preterm ages, and continued follow-up is critical to achieving that goal. The infants included in this study will be followed through at least 2 years chronological age, at which time all of the participants in this study will be administered the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III). The Bayley-III assesses all facets of child development, including motor, cognitive, language, adaptive behavior, and social emotional development (103). Its reliability and validity are well-documented, and it is considered a “gold standard” for developmental assessment through 42 months of age. This comprehensive testing will provide insight as to whether the Premie-Neuro is predictive of developmental outcomes at 2 years chronological age.

Although significant neurodevelopmental deficits can generally be detected by two years of age, it may take longer for more mild delays and disabilities to become apparent. A 2006 study by Salt and colleagues (39) showed that preterm children who were “normal” at 3 years often go on to have significant deficits at school-age, strengthening the argument that routine NICU follow-up should continue beyond the often standard 2 years of age. Recent research has shown that, although their performance may not be in the abnormal range, preterm children generally have below average cognitive, reading, math, and motor skills, performing significantly

worse in these areas than their term peers. This appears to be particularly true for extremely preterm infants born before 28 weeks completed gestation (30, 31, 33, 34, 36, 38, 46-52). These school-age deficits can have a significant impact on school success as well as social emotional and behavioral development. However, there has been little work done researching whether these more mild deficits may be identified earlier in life so that they may be prevented or possibly treated. In order to address this, our work may be extended beyond the 2-year follow-up to assess the study participants at school age and beyond so that we can determine if the Premie-Neuro—or other data collected in this study—may provide an early indication of future mild neurodevelopmental deficits.

Early intervention services after NICU discharge

Studies have shown that general developmental care, such as NIDCAP, has a positive effect on neurodevelopmental outcomes (8, 39, 80). However, more specialized early intervention services, such as physical therapy, have only been shown to significantly affect motor performance if they are performed on infants who are identified as “high risk” using a standardized assessment tool (12, 13). Thus, it seems that early intervention services are beneficial, but for only a subset of preterm infants. There are few data describing factors that influence access to and utilization of early intervention services by preterm infants and their families. In our study, we provided a questionnaire to parents or caregivers at the 3 month follow up, detailing demographic information, parental concerns about their child’s development, and

what sort of early intervention services, if any, the child was receiving. Although these data were not analyzed and presented in this dissertation, they may be studied in the future to determine what factors influence utilization of early intervention services for preterm children and their families. Because we are in the unique situation that this study was conducted near the Kansas-Missouri state line—two states with vastly different criteria for eligibility for their States’ early intervention programs—we may also be able to determine if there are any differences in utilization of services and outcomes in the two states. Finally, for the preterm children who are receiving early intervention services, we may be able to use the data from the questionnaires and the outcome measures used in this study to determine if early intervention services are effective and—if so—if there are neonatal factors that predict positive outcomes.

5.5 Conclusions

In this study, we provide evidence that the Premie-Neuro is a reliable and valid clinical neurological test for preterm infants in the NICU, and that performance on the Premie-Neuro in the NICU is predictive of outcomes at term and 3 months adjusted age. Further, we propose modification of raw score cut-points for classifying an assessment as normal, abnormal, or questionable that may assist the clinician in interpreting Premie-Neuro results and using them to make reliable and valid decisions for NICU services, follow-up, and referral for early intervention services after NICU discharge. This is the first study to validate a standardized assessment tool on preterm infants as young as 29 weeks PMA, and is unique in that

even very medically fragile infants were tested, including those who were ventilator-dependent and who had significant medical complications such as central nervous system injury, retinopathy of prematurity, and necrotizing enterocolitis. Thus, our findings may be generalized to the wide range of preterm infants typically admitted to the NICU. The results of this study expand our understanding of preterm infant development and provide insight on how early neurological signs may be used to predict future neurodevelopmental disability and delay. Future work will further determine the psychometric soundness of the Premie-Neuro, determine how early true developmental disability and delay can be detected, and examine how early intervention services may be best utilized to improve developmental outcomes in preterm children.

References

1. Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Menacker F, Kirmeyer S, et al. Births: final data for 2006: Centers for Disease Control and Prevention. National Vital Statistics System.; 2009 January 7, 2009. Report No.: 57(7).
2. Allen MC. Preterm outcomes research: a critical component of neonatal intensive care. *Mental Retardation and Developmental Disabilities* 2002;8:221-233.
3. Bracewell M, Marlow N. Patterns of motor disability in very preterm children. *Mental Retardation and Developmental Disabilities* 2002;8:241-248.
4. Hintz SR, Kendrick DE, Vohr BR, Poole K, Higgins RD. Changes in neurodevelopmental outcomes at 18 to 22 months' corrected age among infants of less than 25 weeks' gestational age born in 1993-1999. *Pediatrics* 2005;115:1645-1651.
5. Lemons JA, Bauer CR, Oh W, Korones SB, Papile L-A, Stoll BJ, et al. Very low birth weight outcomes of the National Institute of Child Health and Human Development Neonatal Research Network, January 1995 through December 1996. *Pediatrics* 2001;107(1).
6. Kobaly K, Schluchter M, Minich N, Friedman H, Taylor HG, Wilson-Costello D, et al. Outcomes of extremely low birth weight (<1kg) and extremely low gestational age (<28 weeks) infants with bronchopulmonary dysplasia: effects of practice changes in 2000-2003. *Pediatrics* 2008;121:73-81.

7. Robertson CMT, Watt M-J, Yasui Y. Changes in the prevalence of cerebral palsy for children born very prematurely within a population-based program over 30 years. *JAMA* 2007;297(24):2733-2740.
8. Blauw-Hospers CH, Hadders-Algra M. A systematic review of the effects of early intervention on motor development. *Developmental Medicine and Child Neurology* 2005;47(6):421-432.
9. Cameron E, Maehle V, Reid J. The effects of an early physical therapy intervention for very preterm, very low birth weight infants: a randomized controlled clinical trial. *Pediatric Physical Therapy* 2005;17:107-119.
10. Kaaresen PI, Rønning JA, Tunby J, Nordhov SM, Ulvund SE, Dahl LB. A randomized controlled trial of an early intervention program in low birth weight children: outcome at 2 years. *Early Human Development* 2007;84:201-209.
11. Spittle A, Orton J, Doyle L, Boyd R. Early developmental intervention programs post hospital discharge to prevent motor and cognitive impairments in preterm infants. *Cochrane Database of Systematic Reviews* 2007(2):CD005495.
12. Girolami GL, Campbell SK. Efficacy of a neuro-developmental treatment program to improve motor control in infants born prematurely. *Pediatric Physical Therapy* 1994;6:175-184.
13. Lekskulchai R, Cole J. Effect of a developmental program on motor performance in infants born preterm *Australian Journal of Physiotherapy* 2001;47:169-176.

14. Daily D, Ellison PH. The Premie-Neuro: a clinical neurologic examination of premature infants. *Neonatal Network* 2005;24(1):15-22.
15. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet* 2008;371:75-84.
16. Allen M. Assessment of gestational age and neuromaturation. *Mental Retardation and Developmental Disabilities* 2005;11:21-23.
17. Herschkowitz N. Brain development in the fetus, neonate and infant. *Biology of the Neonate* 1988;54:1-19.
18. deGraaf-Peters VB, Hadders-Algra M. Ontogeny of the human central nervous system: What is happening when? *Early Human Development* 2006;82:257-266.
19. Wyatt J. Mechanisms of brain injury in the newborn. *Eye* 2007;21(10):1261-1263.
20. Volpe JJ. Neurobiology of periventricular leukomalacia in the premature infant. *Pediatric Research* 2001;5:553-562.
21. Aylward GP. Neurodevelopmental outcomes of infants born prematurely. *Developmental and Behavioral Pediatrics* 2005;26(6):427-440.
22. Hatzidaki E, Giahnakis E, Maraka S, Korakaki E, Manoura A, Saitakis E, et al. Risk Factors for Periventricular Leukomalacia. *Acta Obstetricia et Gynecologica Scandinavica* 2009 88:110-115.
23. Brouwer A, Groenendaal F, vanHaastert I-L, Rademaker K, Hanlo P, deVries L. Neurodevelopmental outcome of preterm infants with severe intraventricular

hemorrhage and therapy for post-hemorrhagic ventricular dilation. *Journal of Pediatrics* 2008;152:648-654.

24. Broitman E, Ambalavanan N, Higgins RD, Vohr BR, Das A, Bhaskar B, et al. Clinical data predict neurodevelopmental outcome better than head ultrasound in extremely low birth weight infants. *Journal of Pediatrics* 2007;151:500-505.

25. Ramenghi LA, Fumagalli M, Bassi L, Groppo M, Carli AD, Fanaro S, et al. Brain maturation of preterm newborn babies: new insights. *Journal of Pediatric Gastroenterology and Nutrition* 2007;45:S143-S146.

26. Hack M. Survival and neurodevelopmental outcomes of preterm infants. *Journal of Pediatric Gastroenterology and Nutrition* 2007;45:S141-S142.

27. Dyet LE, Kennea N, Counsell SJ, Maalouf EF, Ajayi-Obe M, Duggan PJ, et al. Natural history of brain lesions in extremely preterm infant studies with serial magnetic resonance imaging from birth and neurodevelopmental assessment. *Pediatrics* 2006;118(2):536-548.

28. Schmidt B, Asztalos EV, Roberts RS, Robertson CMT, Sauve RS, Whitfield MF. Impact of bronchopulmonary dysplasia, brain injury, and severe retinopathy on the outcome of extremely low-birth-weight infants at 18 months. *JAMA* 2003;289(9):1124-1129.

29. Burns Y, O'Callaghan M, McDonnell B, Rogers Y. Movement and motor development in ELBW infants at 1 year is related to cognitive and motor abilities at 4 years. *Early Human Development* 2004;80:19-29.

30. Goyen T-A, Lui K, Woods R. Visual-motor, visual-perceptual, and fine motor outcomes in very-low birthweight children at 5 years. *Developmental Medicine and Child Neurology* 1998;40:76-81.
31. Halsey CL, Collin MF, Anderson CL. Extremely low-birth-weight children and their peers. *Archives of Pediatric and Adolescent Medicine* 1996;150:790-794.
32. Latal-Hajnal B, vonSiebenthal K, Kovari H, Bucher HU, Largo RH. Postnatal growth in VLBW infants: significant association with neurodevelopmental outcome. *Journal of Pediatrics* 2003;143:163-170.
33. Marlow N, Roberts BL, Cooke RWI. Motor skills in extremely low birthweight children at the age of 6 years. *Archives of Disease in Childhood* 1989;64:839-847.
34. Marlow N, Roberts L, Cooke R. Outcome at 8 years for children with birth weights of 1250 g or less. *Archives of Disease in Childhood* 1993;68:286-290.
35. McCormick MC. The outcomes of very low birth weight infants: are we asking the right questions? *Pediatrics* 1997;99(6):869-876.
36. Mikkola K, Ritari N, Tommiska V, Salokorpi T, Lehtonen L, Tammela O, et al. Neurodevelopmental outcome at 5 years of age of a national cohort of extremely low birth weight infants who were born in 1996-1997. *Pediatrics* 2005;116(6):1391-1400.
37. Picciolini O, Gianni ML, Vegni C, Fumagalli M, Mosca F. Usefulness of an early neurofunctional assessment in predicting neurodevelopmental outcome in very

low birthweight infants. Archives of Disease in Childhood. Fetal and Neonatal Edition 2006;91:F111-F117.

38. Powls A, Botting N, Cooke RWI, Marlow N. Motor impairment in children 12 to 13 years old with a birthweight of less than 1250 g. Archives of Disease in Childhood 1995;72:F63-F66.

39. Salt A, Redshaw M. Neurodevelopmental follow-up after preterm birth: follow up after two years. Early Human Development 2006;82:185-197.

40. Wood NS, Marlow N, Costeloe K, Gibson AT. Neurologic and developmental disability after extremely preterm birth. The New England Journal of Medicine 2000;343(6):378-384.

41. Anderson PJ, Doyle LW. Cognitive and educational deficits in children born extremely preterm. Seminars in Perinatology 2008;32:51-58.

42. Petrini JR, Dias T, McCormick MC, Massolo ML, Green NS, Escobar GJ. Increased risk of adverse neurological development for late preterm infants. Journal of Pediatrics 2009;154:169-176.

43. Allen M. Neurodevelopmental outcomes of preterm infants. Current Opinions in Neurology 2008;21(2):123-128.

44. O'Shea M. Cerebral Palsy. Seminars in Perinatology 2008;32:35-41.

45. DeGroote I, Vanhaesebrouck P, Bruneel E, Dom L, Durein I, Hasaerts D, et al. Outcomes at 5 years of age in a population-based cohort of extremely preterm infants. Journal of Obstetrics and Gynecology 2007;110:855-864.

46. Arnaud C, Daubisse-Marliac L, White-Koning M, Pierrat V, Larroque B, Grandjean H, et al. Prevalence and associated factors of minor neuromotor dysfunctions at age 5 years in prematurely born children. *Archives of Pediatric and Adolescent Medicine* 2007;161(11):1053-1061.
47. Neubauer A-P, Voss W, Kattner E. Outcome of extremely low birth weight survivors at school age: the influence of perinatal parameters on neurodevelopment. *European Journal of Pediatrics* 2008;167:87-95.
48. Chyi LJ, Lee HC, Hintz SR, Gould JB, Sutcliffe TL. School outcomes of late preterm infants: special needs and challenges for infants born at 32 to 36 weeks gestation. *Journal of Pediatrics* 2008;153(1):25-31.
49. Constable RT, Ment LR, Vohr BR, Kesler SR, Fulbright RK, Lacadie C, et al. Prematurely born children demonstrate white matter microstructural differences at 12 years of age, relative to term control subjects: an investigation of group and gender effects. *Pediatrics* 2008;121:306-316.
50. Davis NM, Ford GW, Anderson PJ, Doyle LW. Developmental coordination disorder at 8 years of age in a regional cohort of extremely-low-birthweight or very preterm infants. *Developmental Medicine and Child Neurology* 2007;49:325-330.
51. Marlow N, Hennesy EM, Bracewell M, Wolke D, Group ES. Motor and executive function at 6 years of age after extremely preterm birth. *Pediatrics* 2007;120(4):793-804.
52. Wocadlo C, Rieger I. Motor impairment and low achievement in very preterm children at eight years of age. *Early Human Development* 2008;84(11):769-776.

53. Goyen T-A, Lui K. Developmental coordination disorder in "apparently normal" schoolchildren born extremely preterm. *Archives of Diseases in Childhood* 2009;94:298-302.
54. O'Shea TM, Kuban KCK, Allred EN, Paneth N, Pagano M, Dammann O, et al. Neonatal cranial ultrasound lesions and developmental delays at 2 years of age among extremely low gestational age children. *Pediatrics* 2008;122(3):e662-e669.
55. Wilson-Costello D, Borawski E, Friedman H, Redline R, Fanaroff AA, Hack M. Perinatal correlates of cerebral palsy and other neurologic impairment among very low birth weight children. *Pediatrics* 1998;102(2):315-322.
56. Rademaker KJ, Uiterwaal CSPM, Beek FJA, vanHaastert IC, Liefink AF, Groenendaal F, et al. Neonatal cranial ultrasound versus MRI and neurodevelopmental outcome at school age in children born preterm *Archives of Disease in Childhood. Fetal and Neonatal Edition* 2005;90:F489-F493.
57. Badr LK, Bookheimer S, Purdy I, Deeb M. Predictors of neurodevelopmental outcome for preterm infants with brain injury: MRI, medical and environmental factors. *Early Human Development* 2009;doi:10.1016/j.earlhumdev.2008.11.005.
58. Constantinou JC, Adamson-Macedo EN, Mirmiran M, Fleisher BE. Movement, imaging and neurobehavioral assessment as predictors of cerebral palsy in preterm infants. *Journal of Perinatology* 2007;27:225-229.
59. Aylward GP, Hatcher RP, Leavitt LA, Rao V, Bauer CR, Brennan MJ, et al. Factors affecting neurobehavioral responses of preterm infants at term conceptional age. *Child Development* 1984;55:1155-1165.

60. Grégoire M-C, Lefebvre F, Glorieux J. Health and developmental outcomes at 18 months in very preterm infants with bronchopulmonary dysplasia. *Pediatrics* 1998;101(5):856-860.
61. Jeng S-F, Hsu C-H, Tsao P-N, Chou H-C, Lee W-T, Kao H-A, et al. Bronchopulmonary dysplasia predicts adverse developmental and clinical outcomes in very-low-birthweight infants. *Developmental Medicine and Child Neurology* 2008;50:51-57.
62. Tyson JE, Parikh NA, Langer J, Green C, Higgins RD. Intensive care for extreme prematurity--moving beyond gestational age. *The New England Journal of Medicine* 2008;358(16):1672-1681.
63. Spinillo A, Montanari L, Gardella B, Roccio M, Stronati M, Fazzi E. Infant sex, obstetric risk factors, and 2-year neurodevelopmental outcome among preterm infants. *Developmental Medicine and Child Neurology* 2009;DOI: 10.1111/j.1469-8749.2009.03273.x:518-525.
64. Clements KM, Barfield WD, Ayadi MF, Wilber N. Preterm birth-associated cost of early intervention services: an analysis by gestational age. *Pediatrics* 2007;119:e866-e874.
65. Majnemer A, Mazer B. Neurologic evaluation of the newborn infant: definition and psychometric properties. *Developmental Medicine and Child Neurology* 1998;40:708-715.

66. Als H, Lester BM, Tronick EZ, Brazelton TB. Toward a research instrument for the assessment of preterm infants' behavior (APIB). New York: Plenum Press; 1982.
67. Snider L, Tremblay S, Limperopoulos C, Majnemer A, Filion F, Johnston C. Construct validity of the Neurobehavioral Assessment of Preterm Infants. *Physical and Occupational Therapy in Pediatrics* 2005;25(3):81-95.
68. Majnemer A, Snider L. A comparison of developmental assessments of the newborn and young infant. *Mental Retardation and Developmental Disabilities* 2005;11:68-73.
69. Als H, Butler S, Kosta S, McAnulty G. The Assessment of Preterm Infants' Behavior (APIB): Furthering the understanding and measurement of neurodevelopmental competence in preterm and full-term infants. *Mental Retardation and Developmental Disabilities* 2005;11:94-102.
70. Korner AF, Brown BW, Dimiceli S, Forrest T, Stevenson DK, Lane NM, et al. Stable individual differences in developmentally changing preterm infants: a replicated study. *Child Development* 1989;60:502-513.
71. Korner AF, Constantinou J, Dimiceli S, Brown BW. Establishing the reliability and developmental validity of a neurobehavioral assessment for preterm infants: a methodological process. *Child Development* 1991;62:1200-1208.
72. Korner AF, Kraemer HC, Reade EP, Forrest T, Dimiceli S. A methodological approach to developing an assessment procedure for testing the neurobehavioral maturity of preterm infants. *Child Development* 1987;58:1478-1487.

73. Campbell SK, Kolobe THA, Osten ET, Lenke M, Girolami GI. Construct validity of the Test of Infant Motor Performance. *Physical Therapy* 1995;75(7):585-596.
74. Campbell SK, Hedeker D. Validity of the Test of Infant Motor Performance for discriminating among infants with varying risk for poor motor outcome. *Journal of Pediatrics* 2001;139:546-551.
75. Rose RU, Westcott SL. Responsiveness of the Test of Infant Motor Performance (TIMP) in infants born preterm. *Pediatric Physical Therapy* 2005;17:219-224.
76. Lester BM, Tronick EZ. History and description of the Neonatal Intensive Care Unit Network Neurobehavioral Scale. *Pediatrics* 2004;113(3):634-640.
77. Salisbury AL, Fallone MD, Lester B. Neurobehavioral assessment from fetus to infant: the NICU Network Neurobehavioral Scale and the Fetal Neurobehavior Coding Scale. *Mental Retardation and Developmental Disabilities Research Reviews* 2005;11:14-20.
78. Daily D, Cannon S. The Premie-Neuro examination instruction manual. In; 2005.
79. Cannon S, Weatherstone KB, Daily D. Early neurodevelopmental assessment in the newborn intensive care unit (abstract). *Pediatric Research* 2002;51(4):401.
80. Als H, Duffy FH, McAnulty G, Rivkin MJ, Vajapeyam S, Mulkern RV, et al. Early experience alters brain function and structure. *Pediatrics* 2004;113(4):846-857.

81. Heathcock JC, Lobo M, Galloway JCC. Movement training advances the emergence of reaching in infants born at less than 33 weeks of gestational age: a randomized clinical trial. *Pediatric Physical Therapy* 2008;88(3):310-322.
82. Liu W, Laudert S, Perkins B, MacMillan-York E, Martin S, Graven S. The development of potentially better practices to support the neurodevelopment of infants in the NICU. *Journal of Perinatology* 2007;27:S48-S74.
83. Payton OD. Research: the validation of clinical practice. 3rd ed. Philadelphia: F.A. Davis Company; 1994.
84. Wang H-H, Liao H-F, Hsieh C-L. Reliability, sensitivity to change, and responsiveness of the Peabody Developmental Motor Scales--Second Edition for children with cerebral palsy. *Physical Therapy* 2006;86(10):1351-1359.
85. Brazy JE, Goldstein RF, Oehler JM, Gustafson KE, Thompson RJ. Nursery Neurobiologic Risk Score: levels of risk and relationships with nonmedical factors. *Journal of Developmental and Behavioral Pediatrics* 1993;14:375-380.
86. Brazy JE, Eckerman CO, Oehler JM, Goldstein RF, O'Ran AM. Nursery Neurobiologic Risk Score: Important factors in predicting outcome in very low birth weight infants. *Journal of Pediatrics* 1991;118:783-792.
87. Lefebvre F, Gregoire M-C, Dubois J, Glorieux J. Nursery Neurobiologic Risk Score and outcome at 18 months. *Acta Paediatrica* 1998;87:751-757.
88. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 1977;33(1):159-174.

89. Deschenes G, Gosselin J, Couture M, Lachance C. Interobserver reliability of the Amiel-Tison Neurological Assessment at Term. *Pediatric Neurology* 2004;30(3):190-194.
90. Ellison PH. *The Infanib: A Reliable Method for the Neuromotor Assessment of Infants*. San Antonio, TX: Harcourt Assessment, Inc.; 1994.
91. Morgan AM, Koch V, Lee V, Aldag J. Neonatal Neurobehavioral Examination: a new instrument for quantitative analysis of neonatal neurological status. *Physical Therapy* 1988;68(9):1352-1358.
92. Campbell S. Test-retest reliability of the Test of Infant Motor Performance. *Pediatric Physical Therapy* 1999;11:60-66.
93. MacDorman MF, Callaghan W, Mathews TJ, Hoyert DL, Kochanek KD. Trends in preterm-related infant mortality by race and ethnicity: United States, 1999-2004. In: *Statistics NCfH*, editor.; 2007.
94. Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Menacker F, Munson ML. *Births: final data for 2002: Centers for Disease Control and Prevention. National Vital Statistics System.*; 2002.
95. Sheridan-Pereira M, Ellison PH, Helgeson V. The construction of a scored neonatal neurological examination for assessment of neurological integrity in full-term neonates. *Journal of Developmental and Behavioral Pediatrics* 1991;12:25-30.
96. Piper M, Darrah J. *Motor Assessment of the Developing Infant*. Philadelphia: WB Saunders; 1994.

97. Darrah J, Piper M, Watt M-J. Assessment of gross motor skills of at-risk infants: predictive validity of the Alberta Infant Motor Scale. *Developmental Medicine and Child Neurology* 1998;40:485-491.
98. Portney LG, Watkins MP. *Foundations of clinical research: Applications to practice*. 2nd ed. Upper Saddle River, NJ: Prentice Hall; 2000.
99. Spittle A, Doyle LW, Boyd R. A systematic review of the clinimetric properties of neuromotor assessments for preterm infants during the first year of life. *Developmental Medicine and Child Neurology* 2008;50(4):254-266.
100. Aylward GP. Methodological issues in outcome studies of at-risk infants. *Journal of Pediatric Psychology* 2002;27(1):37-45.
101. Hines RB, Minde K, Marton P, Trehub S. Behavioral development of premature infants: an ethological approach. *Developmental Medicine and Child Neurology* 1980;22:623-632.
102. Paludetto R, Rinaldi P, Mansi G, Andolfi M, Giudice GD. Early behavioral development of preterm infants. *Developmental Medicine and Child Neurology* 1984;26:347-352.
103. Bayley N. *Bayley Scales of Infant and Toddler Development, Third Edition*. San Antonio, Texas: Harcourt Assessment, Inc.; 2005.